Regulating with ribonucleases in *Streptococcus pyogenes*

Dissertation

zur Erlangung des akademischen Grades

Doctor of Philosophy (Ph.D.)

eingereicht an der

Lebenswissenschaftlichen Fakultät

der Humboldt-Universität zu Berlin

von

Laura Broglia, M.Sc.

Präsidentin der Humboldt-Universität zu Berlin

Prof. Dr.-ing Dr. Sabine Kunst

Dekan der Lebenswissenschaftliche Fakultät

der Humboldt-Universität zu Berlin

Prof. Dr. Bernhard Grimm

Gutachter/innen

1. Prof. Dr. Emmanuelle Charpentier
2. Prof. Dr. Markus Landthaler
3. Prof. Dr. Petra Dersch

Tag der mündlichen Prüfung: 5. June 2020

"You are lost the moment you know what the result will be."

Juan Gris

1887-1927

Acknowledgements

I would like to thank the people who have made this thesis possible and without whom, I would not have been able to reach this achievement.

I am grateful to my supervisor Prof. Emmanuelle Charpentier for accepting me at first as a master student and then as a PhD candidate in her group(s). Thank you for these multiple opportunities and the trust, for those thrilling moments of paper submissions and for the one thing in common between my PhD thesis and yours!

I would like to thank the members of my thesis committee Prof. Markus Landthaler, Prof. Petra Dersch and Prof. Pascale Romby for the yearly TAC meetings, for their time and suggestions. I would also like to thank Prof. Kürşad Turgay for the corrections on the thesis and manuscripts and for the enriching discussions on the microbe world. By now, I hope I have convinced you that RNases perform RNA *decay*.

My deepest gratitude goes to Anaïs. You have been a wonderful co-supervisor. Thank you for all you have been teaching me, for your support in all these years, for your dedication, for your precious advices, for our endless email and phone conversations, for always finding the time to discuss (and not only about science) and to correct the multiple versions of reports, presentations, manuscripts, thesis etc. I feel that I scientifically and personally grew up so much with you!

I am grateful to my best co-author and gel loading partner (as you said) Anne-Laure. Thank you for sharing with me all your knowledge about our favorite enzymes (*i.e.* ribonucleases), cloning and knock out strategies. Listening to you talking about science was always so inspiring! Thank you for your support and constantly believing in me.

Special thanks to all the people whom I have met in this journey, all the members of REGI, RIIB and MPUSP including the ones I will not explicitly mention. Andrés, I have been so lucky to have shared most of the PhD with you. Although I was always complaining, thank you for your repetitive jokes, repetitive stories and repetitive jokes and stories. Thank you for your positivity, for being a life-teacher, for always having the right advise (*i.e.* Google it) and being there for me in the most difficult moments. Thank you for our scientific discussions on *speB* and *sagA*, for the "big gel" assembly sessions and for the Inkscape trouble shootings (that always ended up crashing my computer). Thank you for all the breakfasts, lunches, dinners, parties and for being a real friend also outside the lab. There are only two things I will never thank you for: (i) your mocking my Italian accent and (ii) your advices on ear drops and medications. Of course, thank you Signe for being always interested in my life and in the developments of my PhD. Lina, I still remember

when I arrived in Berlin and you welcomed me at the train station. It was cold and grey (standard winter day in Berlin); I did not have a house, I was lost and sad, but with your Latin-American energy you made me feel already like at home. Thank you for taking care of me, for the Sunday Latino brunches, for the unforgettable parties, for our time in "Boooston", for the unexpected and amazing trip to Colombia and of course for our shared love for chips. You have been a great friend and such a good scientific example for me! Eric, thank you for the Sunday mornings with coffee and *arepas* and for sharing your knowledge on the human microbiota. Frank, thank you for the several coffee, cigarette and soup breaks talking about science and life (maybe more about life). I would like to thank *this person* for the thesis writing together, for the long nights in the lab and the long weekends in the office writing and crying and for the cinema and the Christmas market nights. Thank you for the help with German and your always rational advises. Majda, thank you for allowing me to watch Grey's Anatomy (yes, I admit it I like Grey's Anatomy), which was one of the only reasons to leave the lab. Thank you for the fun during the intense working days in the hot-lab and for always listening to my problems. Ines, thank you for all the advises, for the discussions, for the fun, for the shared love for Grey's anatomy and of course for the German version of the thesis summary. Hi Shi Pey, Hi Shi Pey, Hi Shi Pey! Thank you for letting me succeed in the most challenging competition of my PhD. Hagen, thank you for including me in the *bubu* community as *lala* (although I know I was never really part of it). James, thank you for the few months of office sharing, for your interest in the SpeB story and of course for being such a good organizer of lab events. Thibaud, thank you for always having the right questions in the meetings and the right answers in the discussions. Thank you for the thousands or more RNA structure predictions! Karin, you are the perfect technician and working with you was so productive. Thank you for your help, your interest in the projects, your tips in the lab and for always reminding me of the semester fees! Stefan, thank you for your craziness and for always sharing the trash and fatty food you were storing in the office. Victoria, you are amazing! Thank you for being such a good friend, for the motivation, energy and the good vibes you brought in the lab. Thank you for your support in the long nights spent taking time points or in the hot-lab performing Northern blot analyses or in our favorite bar in Berlin *La cantina*. Thank you for the Pilates lessons and for our weekend plans. Solange, you have been a great student! Thank you for your dedication and effort in trying to solve the *speB*-RNase Y mystery. Vanessa, thank you for your support and empathy and for always being there in the worst of the moments. You are a real science-driven person and I have always learned something from you. Thank you for making me discover a different Berlin, the one of our favorite restaurants, bars and shops, the one we nimbly and incautiously cross with our bikes. Thank you for the joint effort in trying to improve our lifestyle (*e.g.* sport, healthy food, savings…). I do not know how much we have accomplished, but the planning and the discussion were great! Matteo,

lo so che avresti voluto un capitolo dedicato esclusivamente al conte Ugolini, ma ti devi accontentare di un paragrafo. How could I start thanking the person that saved this thesis? Matteo, without you, this thesis would not physically exist! Thank you for the support and help especially in these last months and for standing my stress and anxiety. Thank you for the stylish chapter numbers and for the miracle of the appendix tables. Thank you for your perfect suggestions on almost everything, from the right fonts and colors to what shoes and dress to buy. Your aspiration to perfection and high-quality standards are now with me every time I start writing or drawing. You are an excellent scientist and you have been a great inspiration for me and not only scientifically. Thank you for the nice time we have spent together in chic restaurants, classy bars and exclusive Caribbean islands.

I would like also to thank my office colleagues Daniel, Nina, Sandra and Thomas for coping with my crazy and stressful moments, with my spreading in the office and with my obsession of turning the light off. Thank you for the encouragement and the office breaks. Thank you to Marlene, Dior, Mirshat and all the PhDs for the funny moments, the shared organization of retreats and happy events and the PhD breakfasts. Rina, Lya and Lisa, thank you for your help and for the international dinners. Tim, thank you for always having good comments on my numerous presentations. Gisela and Nina thank your support and advices.

I would like to thank my multiple flat mates who in all these years they probably thought I was a ghost, thank you for your understanding. Simone and Giselle, thank you for your interest in my studies, the dinners together and the beer tasting sessions.

Many thanks to all my friends at home, the ones from my little town, from school, from university…thank you for always encouraging me, for your calls and messages. Thank you for the nice time spent together in my short trips to Italy. You helped me to go through all of this.

Daniele, where do I start? Thank you for always being there for me, for accepting and supporting my decisions, for the unlimited patience with which you face my complaints, anxiety and crazy ideas. Thank you for believing in this, even if it was not easy! Thank you for your trips to Berlin, for the awesome Italian weekends and for all the other trips together; all this time spent with you has been always a source of motivation and joy.

Finally, my deepest gratitude goes to my family. Ai miei genitori, mia sorella e ai miei piccoli nipoti grazie per avermi lasciato partire e prendere la mia strada anche se non è stato facile, lo so. Grazie perché anche se lontana, ogni volta mi fate sentire come se non me ne fossi mai andata. Grazie per il vostro sostegno incondizionato e per credere in me sempre. Grazie per i vostri viaggi a Berlino o negli aeroporti ad aspettarmi, per le videochiamate serali trascorse ad ascoltare i miei problemi. Grazie perché senza il vostro supporto non sarei mai riuscita a raggiungere questo traguardo!

Table of content

List of figures

List of tables

Summary

Bacteria have developed a plethora of strategies to cope with constantly changing environmental conditions, including post-transcriptional regulatory mechanisms. With this regard, regulation of gene expression can be achieved by either the rapid removal or stabilization of RNA molecules by ribonucleases (RNases). RNases properties and functions have been mainly studied in the two model organisms *Escherichia coli* and *Bacillus subtilis* and their investigation often relied on the characterization of a limited number of substrates. Although these studies are highly informative, they do not provide a global picture of all the RNase targets in the bacterial cell. In addition, studies in non-model microorganisms – including pathogens – have revealed that distinct RNase orthologs exhibit species-specific effects on gene expression, bacterial physiology and strategies of target recognition, indicating that our understanding of the RNA degradation machinery is not yet complete.

The aim of this thesis was to investigate the features and functions of endoRNase Y from the strict human pathogen *Streptococcus pyogenes*, whose RNA degradation machinery is poorly studied. To gain insight into the role and specificity of this RNase, we identified RNase Y cleavage positions (*i.e.* targetome) genome-wide by RNA sequencing. Next, to investigate the RNA degradation pathway depending on RNase Y, we compared the RNase Y targetome with the ones of the three 3'-to-5' exoribonuclease (exoRNases) present in this bacterium, namely PNPase, YhaM and RNase R. Finally, to dissect the requirements for RNase Y processing *in vivo* and to decipher the role of RNase Y in virulence gene regulation, we studied the impact of RNase Y on *speB* mRNA, that encodes a major virulence factor.

This study reveals that RNase Y preferentially cleaves transcripts downstream of a guanosine and for the first time we were able to show that the presence of a guanosine residue is essential for the processing of *speB* mRNA, *in vivo*. Although RNase Y cleaves the *speB* mRNA, our data underpin a model in which RNase Y-mediated regulation of *speB* expression occurs at the transcriptional level. Using the targetome comparative approach, we demonstrated that RNase Y initiates RNA decay in *S. pyogenes* and that the transcript 3' ends produced by this RNase are usually further trimmed by the 3'-to-5' exoRNases. In particular, RNase Y acts mainly in concert with PNPase in RNA decay, through the generation and subsequent degradation of decay intermediate fragments.

The comprehensive analysis of RNase Y processing positions gives insights into the specific features of RNase Y in *S. pyogenes*, highlighting the importance of expanding the investigation of RNase Y orthologs to poorly characterized organisms. Overall, these findings increase our understanding of RNase Y functionality and RNA degradation in Gram-positive bacteria.

Zusammenfassung

Bakterien haben eine Vielzahl von Strategien entwickelt, um sich an ständig wechselnde Umweltbedingungen anzupassen, dazu gehören unter anderem post-transkriptionelle regulatorische Mechanismen. Genexpression kann entweder durch einen zügigen Abbau oder die Stabilisierung von RNA durch Ribonukleasen (RNasen) reguliert werden. Eigenschaften und Funktion von RNasen wurden bisher hauptsächlich in *Escherichia coli* und *Bacillus subtilis* untersucht, wobei diese Charakterisierung oft auf einer limitierten Anzahl von Substraten beruht. Obwohl diese Studien informativ sind, können sie kein umfassendes Bild aller RNase-Substrate in Bakterien widerspiegeln. Weiterhin haben Untersuchungen in Nicht-Modell Organismen, inklusive Pathogenen gezeigt, dass bestimmte orthologe RNasen spezies-spezifische Einflüsse auf Genexpression, bakterielle Physiologie und Strategien der Substraterkennung haben. Dies unterstreicht die Tatsache, dass unser Verständnis des RNA Abbaus bei weitem nicht vollständig ist.

Ziel dieser Arbeit ist es die Eigenschaften und Funktionen der endoRNase Y des Humanpathogen Bakteriums *Streptococcus pyogenes* zu untersuchen. Um einen Überblick über die Rolle und Spezifität dieser RNase zu gewinnen, wurden deren genomweiten Schnittpositionen ("targetome") mit Hilfe von RNA-Sequenzierung identifiziert. Dieses Ergebnis wurde mit dem "targetome" der drei 3'-nach-5' Exoribonukleasen (exoRNasen) PNPase, YhaM und RNaseR verglichen, um den RNA-Abbauweg von RNase Y zu untersuchen. Anforderungen für RNase Y Aktivität *in vivo* und deren Rolle in der Regulation von Virulenz-Genen wurde anhand des Einflusses von RNAse Y auf die *speB* mRNA, welche einen wichtigen Virluenz-Faktor codiert, untersucht. Wir konnten in dieser Arbeit zeigen, dass RNase Y RNA-Substrate präferentiell nach einem Guanosin schneidet und dass dieses Nukleosid essentiell für die Prozessierung der *speB* mRNA *in vivo* ist. Obwohl RNase Y die *speB* mRNA schneidet, unterstützen die Daten ein Modell nach dem RNase Y die Expression von *speB* auf der Ebene der Transkription reguliert. Mit Hilfe des "targetome" Vergleichs konnten wir zeigen, dass RNase Y den RNA Abbau in *S. pyogenes* initiiert und dass die dabei produzierten 3' Enden der RNA weiter von 3'-nach-5' exoRNasen getrimmt werden. Der RNA Abbau erfolgt dabei vor allem durch die Generierung und den darauffolgenden Abbau von intermediären Fragmenten durch das Zusammenspiel von RNase Y mit PNPase.

Die umfassende Analyse der RNase Y Prozessierungsstellen gewährt Einblicke in spezifische Eigenschaften von RNase Y in *S. pyogenes*. Dies betont die Wichtigkeit der Untersuchung weiterer RNase Y Orthologe in weniger charakterisierten Bakterien. Zusammenfassend erweitern diese Erkenntnisse unser Verständnis der Funktionalität von RNase Y und des RNA Abbaus in Gram-positiven Bakterien.

Introduction

1.1 Post-transcriptional regulation of gene expression in bacteria

Bacteria need to constantly cope with changing environmental conditions in order to successfully acquire nutrients and proliferate. Such a complex adaptation is achieved by precisely adjusting gene expression according to the external stimuli and environmental cues. For a long time, bacterial gene expression regulation was thought to occur entirely at the transcriptional level, through the regulated activity of transcriptional factors. In 1961, Jacob and Monod in the description of the "operon model", formulated the hypothesis that the repressor molecule, responsible for inhibiting the operon expression, was a ribonucleic acid (RNA) molecule acting either at the deoxyribonucleic acid (DNA) or at the messenger RNA (mRNA) level[1]. In the latter case, the operon negative regulation would occur post-transcriptionally[1]. While, this mechanism of gene expression regulation was initially neglected with the discovery of the first protein acting as transcriptional repressors (*i.e.* Lac and Lambda repressors)[2,3], research over the past few decades has revealed that bacterial gene expression is far more complex than what was thought and does not uniquely depend on transcriptional factors. It is now, indeed, well established that bacteria also extensively control the expression of genes at the post-transcriptional level through an intricate network of regulatory small RNAs (sRNAs), RNA binding proteins and ribonucleases (RNases).

1.1.1 Regulatory RNAs

RNA molecules are typically divided into: (i) mRNAs involved in transferring the genomic information from DNA to proteins, (ii) ribosomal RNAs (rRNAs) representing the major structural component of ribosomes, (iii) transfer RNAs (tRNAs) responsible for decoding the mRNAs into proteins and (iv) sRNAs, which exert regulatory roles and usually do not encode

proteins. The first RNA described to function as regulatory molecule was an antisense RNA (asRNA), encoded within the *Escherichia coli* plasmid ColE1 and named RNAI[4]. RNAI inhibits the plasmid replication by binding to RNAII, which hybridizes to the DNA template and serves as a primer for the DNA polymerase[5].

sRNAs play a key role in the control of gene expression at the post-transcriptional level and are involved in a variety of physiological processes in bacteria, such as virulence, stress responses and community behavior[6]. sRNAs can be divided in different categories depending on their mode of action and on the nature of their targets: (i) asRNAs that are *cis*-encoded on the opposite DNA strand compared to the target RNA, (ii) riboswitches that are encoded within the 5' untranslated region (5' UTR) of the target RNA that they regulate in *cis*, (iii) *trans*-acting sRNAs encoded on a different locus than the target mRNA, (iv) sRNAs interacting with proteins instead of RNA molecules and (v) clustered regularly interspaced palindromic repeats (CRISPR) associated sRNAs, which interact with DNA molecules[6,7].

asRNAs

asRNAs control gene expression by base-paring the target RNAs. Since asRNAs are encoded in *cis* on the opposite DNA strand compared to their target RNAs, the two RNAs share perfect sequence complementarity. At first, asRNAs were mainly thought to be encoded in mobile genetic elements (*e.g.* plasmids, transposons and bacteriophages) where they are involved in the control of plasmid replication or bacteriophage development[8,9].

Plasmids often encode toxin-antitoxin (TA) systems consisting of stable toxin and an unstable antitoxin, which neutralizes the toxin activity. These systems ensure plasmid stability by post-segregational killing since the more stable toxin will kill the cells that have lost the plasmid. In type I TA systems, the antitoxin is an asRNA that base-pairs with the mRNA encoding the toxin and inhibits its expression[10,11]. The best characterized type I TA system is the host killing (*hok*)/suppression of killing (*sok*) system encoded in the *E. coli* plasmid R1[12]. The antitoxin RNA Sok interacts with the Shine-Dalgarno (SD) sequence of the open reading frame (ORF) located just upstream of the toxin gene *hok* and indirectly inhibits *hok* mRNA translation. The duplex formed by the two RNAs is then recognized and subsequently cleaved by the double-stranded RNA (dsRNA) specific endoribonuclease (endoRNase) III (RNase III)[13,14].

With the development of highly sensitive technologies to study bacterial transcriptomes, a large amount of asRNA was also identified to originate from antisense transcription of bacterial genomes[15–18]. asRNAs, usually between 100 and 300 nucleotides (nt) long, can be classified based on whether they derive from antisense transcription of 5' UTRs, 3' UTRs or from ORFs[19]. Longer

asRNAs have been also reported, for example in *Listeria monocytogenes*, where some asRNAs were detected that span over several ORFs[20].

Different mechanisms are exploited by asRNAs to control gene expression at the level of transcription, RNA stability and translation. In the first scenario, asRNAs can affect transcription by directly interfering with the transcription machinery of the sense strand or by transcription attenuation, which affects the formation of a termination structure downstream of the asRNA-RNA duplex[21]. At the post-transcriptional level, the asRNA binding to the target mRNA leads to occlusion of the ribosome binding site (RBS) and to subsequent inhibition of the translation initiation. asRNA-target mRNA interaction can also induce processing by RNases. For instance, the asRNA GadY, base-pairs with the *gadXW* bicistronic mRNA and induces cleavage by RNase III, which processes the duplex RNA, causing *gadW* mRNA stabilization[22–24].

Long asRNAs (lasRNAs) encoded in the antisense DNA strand were shown not only to act as mRNAs (encoding for proteins), but also to exert regulatory functions by inhibiting the expression of the genes expressed in the sense DNA strand[21]. The outcome of this genomic arrangement is that the lasRNA negatively regulates the sense ORFs while ensuring the expression of the antisense (*i.e.* divergently transcribed) ORFs. This distinctive locus arrangement was named "excludon", as it exclusively allows the expression of one of the two transcriptional units[21]. Recently, the concept of not-contiguous operon, consisting of genes co-transcribed from the sense DNA strand and interspaced by a gene divergently transcribed, was proposed in *Staphylococcus aureus*[25]. This organization would allow the coordinated regulation of genes encoded in the same operon, by RNase III mediated-processing of the overlapping region, which in turns affects mRNA translation and/or stability[25].

Trans-acting sRNAs

trans-acting sRNAs are encoded at a distal site in the genome compared to the target mRNA. Hence, there is not perfect sequence complementarity between the sRNAs and the mRNA target. While originally *trans*-acting sRNAs that harbor their own promoter and terminator were only identified within intergenic regions[26,27], it is now clear that these sRNAs can also originate from 5' and 3' UTRs of mRNAs[28–30]. In addition, fragments derived from degradation of mRNAs and tRNAs can also act as regulatory RNAs[31,32].

trans-acting sRNAs mainly regulate the target mRNA by sequestering the RBS and thereby blocking the 30S ribosome recruitment. The subsequent lack of ribosomes on the mRNA exposes the mRNA to the activity of RNases, which therefore degrade the targeted mRNA[6,33]. Alternatively, *trans*-acting sRNAs can positively regulate gene expression by promoting the mRNA translation.

In this case, binding of the sRNA to the target mRNA triggers a structural change that is responsible for the liberation of the sequestrated RBS, which otherwise is inaccessible to the ribosome[34].

trans-acting sRNAs do not only act at the mRNA 5' UTRs, but can also target the mRNA internally, *i.e.* in the coding DNA sequence (CDS). MicC, the first sRNA identified to interact with the target mRNA in the CDS, was shown to induce the degradation of the *ompD* mRNA by recruiting an endoRNase (RNase E) and promoting degradation initiation of the target in *Salmonella typhimurium*[35,36]. sRNAs can also prevent mRNA degradation and therefore causing mRNA stabilization. For example, in *Bacillus subtilis* the RoxS sRNA interacts with the 5' end region of *yflS* mRNA thereby blocking the activity of RNase J1, an exoribonuclease (exoRNase) that degrades RNA molecules from the 5' extremity[37]. The interplay between sRNAs and RNases will be further discussed in section 1.1.6.

Finally, sRNAs can also act in *trans* by interacting with others sRNAs and therefore indirectly regulating RNA targets. For instance, the SroC sRNA, consisting of an RNA fragment produced during RNA degradation of the *gltIJKL* mRNA (coding for an amino acid transporter) functions as a sponge of the GcvB sRNA[38]. Interestingly, it has also been shown that RNA fragments produced by processing of the pre-tRNA can act as RNA sponges to counteract sRNA activity[39].

Riboswitches

RNA-dependent regulation at the post-transcriptional level also relies on riboswitches that consist of RNA structures found in the 5' UTR of mRNAs. These elements are able to sense different type of metabolites (*e.g.* vitamins cofactor, amino acids) or environmental conditions (*e.g.* temperature)[40–42]. As they regulate the expression of their own mRNA, riboswitches are defined as *cis*-acting regulatory elements. The binding of a specific metabolite or the sensing of a certain stimulus induces a conformational change that modulates the expression of the downstream located CDS. Riboswitches affect gene expression through different mechanisms, including inhibition or activation of the mRNA translation and transcriptional termination or antitermination.

A different type of 5' regulatory element largely used by Gram-positive bacteria is the so-called T-box, which senses non-aminoacylated tRNAs[43,44]. The interaction with the uncharged tRNAs leads to the formation of an antitermination structure, which promotes transcription of the downstream ORF encoding a specific aminoacyl-tRNA synthetase.

Interestingly, some riboswitches can also function as *trans*-acting regulators, as shown in *L. monocytogenes*[28]. The premature terminated transcript of the S-adenosyl-L-methionine riboswitch binds to the 5' UTR of *prfA* mRNA, encoding the major virulence transcriptional regulator, and blocks its translation[28].

sRNA sequestering proteins

Some others sRNAs act in *trans* by interacting with proteins in order to sequester them or inhibit their function. This distinct group of sRNAs includes the CsrB sRNA that binds the carbon storage regulator CsrA, consisting of a protein that negatively regulates several target mRNAs through translation inhibition[45].

Another member of this class is the 6S sRNA, which was shown to associate with the RNA polymerase (RNAP) complex and modulate its activity[46–48]. Bacterial RNAP consists of a multicomplex, which includes the factors involved in transcription and the sigma specificity factor (σ) responsible for the recognition of the promoter regions. Bacteria display a housekeeping σ factor, known as σ^{70} or σ^{A}, which is required in all growth stages and also different alternative σ factors, which are activated under specific environmental conditions[49]. During stationary phase of growth, the 6S sRNA interacts with the RNAP complex through the σ^{70} factor, mimicking the interaction of RNAP with the promoter regions, and therefore prevents the RNAP binding to the DNA. The outcome is the repression of genes harboring a σ^{70}-dependent promoter[47,48,50].

sRNAs interacting with proteins are also largely used in type III TA systems, as exemplified by the *toxI*/*toxN* locus of *Erwinia carotovora*[51]. The RNA antitoxin (ToxI) is produced as a precursor long RNA, which comprises an array of repeat sequences. This precursor undergoes maturation though the processing by the endoRNase toxin (ToxN), resulting in the production of ToxI monomers. Three ToxI monomers form a complex with three ToxN proteins, leading to the toxin inhibition[10,51].

CRISPR RNAs

Lastly, the sRNAs involved in the bacterial adaptive clustered regularly interspaced short palindromic repeats (CRISPR) – CRISPR associated proteins (CRISPR-Cas) system constitute another class of sRNAs. CRISPR-Cas systems confer protection against invading foreign nucleic acids, including plasmids and phages. This adaptive defense mechanism relies on the integration of DNA fragments, deriving from foreign nucleic acids, into the genomic CRISPR array locus, allowing to keep memory of the infections[52]. During a second foreign nucleic acid invasion, the CRISPR RNAs (crRNAs) are produced, from endoribonucleolytic processing of the transcribed

CRISPR array (pre-crRNA). The defense mechanism relies on the recognition of the invading DNAs by the crRNAs, which act in *trans* by base-pairing the target DNAs. The crRNAs guide the Cas effector protein(s) to the extrachromosomal DNAs, leading to their subsequent degradation[52].

The crRNAs are generated as a precursor RNA molecule (pre-crRNA) that undergoes a maturation process in order to generate the active crRNAs. In the type II CRISPR-Cas system, the crRNAs biogenesis requires the activity of the *trans*-acting crRNA (tracrRNA), a sRNA that functions in *trans* through base-pairing the pre-crRNAs and promoting the pre-crRNAs processing by RNase III[53,54].

CRISPR-Cas can also be involved in the regulation of endogenous gene expression[55–58]. Recently, it has been shown in *Francisella novicida* that the expression of two bacterial lipoproteins depends on a complex formed by tracrRNA, another small CRISPR-Cas associated RNA (named scaRNA) and the Cas9 nuclease[59]. In this complex, the scaRNA guides Cas9 to interact in proximity of the promoter region of the target genes, promoting transcription inhibition[59].

1.1.2 RNA binding proteins

RNA binding proteins can exert structural roles in large complexes (*e.g.* ribosomes) or can function as regulatory factors involved in transcription, translation and RNA degradation[60].

RNA chaperones

Hfq is an RNA chaperone crucial in the sRNA-mediated regulation, especially in enterobacteria, and it was initially identified as a host factor for the replication of the Qβ phage in *E. coli*. In enterobacteria, Hfq assists *trans*-acting sRNAs in the recognition and regulation of the target mRNAs. Hfq binds both the *trans*-acting sRNA and the target mRNA, promoting the base-pairing and thereby facilitating the formation of the sRNA-mRNA complex[6,61]. Whereas the interaction with the *trans*-acting sRNA occurs though an uracil (U)-rich region at the sRNA 3' end[62], the binding to the mRNA is mediated by an adenosine (A)-rich region at the 5' end region of the target mRNA[63,64]. The complex of Hfq-sRNA is able to recruit RNase E, which is responsible for the degradation of the target mRNA[35,65,66]. However, it is still unclear whether RNase E recruitment is uniquely mediated by the sRNA or by Hfq[60].

Hfq likely plays a different role in Gram-positive bacteria. It was shown indeed to not be required for the sRNA-mediated regulation of gene expression in *S. aureus* and *B. subtilis* and it is completely absent in other bacteria like *Streptococcus pyogenes*[67–69]. The only sRNA identified in Gram-positives that requires Hfq to base-pair its target mRNA is LhrA from *L. monocytogenes*[70]. Based on the Hfq-limited function in Gram-positive bacteria and the fact that some sRNAs in enterobacteria

are Hfq-independent, it is likely that other RNA binding proteins can replace Hfq function. In *B. subtilis*, the Fur-regulated basic protein B (FbpB) was shown to participate in the sRNA FsrA-mediated regulation of gene expression. It was proposed that FbpB may act as an RNA chaperone, however the FbpB ability to bind RNA has not been proven yet[71]. The RNA-binding proteins RocC from *Legionella pneumophila*[72] and ProQ from *Salmonella enterica*[73] were recently identified to be involved in sRNA stabilization and to enhance the sRNA-mRNA duplex formation[72,74]. Oppositely to Hfq, which recognizes sequence motifs, both ProQ and RocC recognize stem-loop structures, such as terminator stem-loops[75,76].

RNA helicases

Bacterial helicases are also key modulators of gene expression at the post-transcriptional level as they affect mRNA degradation and translation[77,78]. Helicases are divided in 6 superfamilies and most of the bacterial helicases belong to the superfamily 2 and to the DEAD-box class. The transcription termination factor Rho represents an exception, as it belongs to the superfamily 5 of helicases[77,79].

The Rho protein consists of a hexameric RNA-dependent ATPase which slides, in 5'-to-3' direction, along a nascent RNA molecule and promotes transcription termination by unwinding of the RNA-DNA duplex[79]. Upon Rho-mediated transcription termination in *E. coli*, the 3' end region of the transcripts is processed by RNases, in order to produce the mature transcript 3' ends[80].

The DEAD-box class of helicases plays a key role in RNA degradation by unfolding RNA secondary structures and therefore generating single-stranded RNA (ssRNA), which can be targeted by ssRNA-specific RNases. In *E. coli*, the RhlB helicase is found associated to the degradosome, which is a multi-enzymatic complex involved in RNA degradation (see section 1.1.5)[81,82]. Other members of the DEAD-box class, such as CsdA, are involved in the regulation of the translation process. In *E. coli*, CsdA was shown to stimulate protein synthesis by unfolding RNA structures, which negatively affect the translation process[83].

Gram-positive bacteria harbour a different set of helicases. In *B. subtilis*, four DEAD-box RNA helicases were identified: CshA, CshB, YmfL and DeaD[84]. Deletion of the genes encoding CshA, CshB or YmfL resulted in impaired growth at low temperatures[85–87]. The DEAD-box helicase CshA, similarly to RhlB, is highly involved in RNA degradation by generating ssRNA regions more susceptible to the exoRNases unable to degrade structured RNAs. As described for RlhB, CshA was shown to interact with RNases and likely be part of a degradosome-like complex in Gram-positive bacteria[88] (see section 1.1.5). Of note, in *S. aureus*, CshA promotes the degradation of the *agr* mRNA, encoding the multiple factors of the *agr* quorum sensing system, including the

AgrA transcriptional regulator[89]. AgrA not only regulates the transcription of the *agr* operon itself, but also the expression of the sRNA RNAIII, a key regulator of virulence[90]. Overall, CshA-mediated repression of *agrA* expression results in downregulation of RNAIII, which in turn leads to increase in biofilm formation and reduction of haemolysin secretion in *S. aureus*[89].

1.1.3 Ribonucleases

RNases are enzymes that catalyze the scission of a phosphodiester bond in an RNA molecule. The RNA susceptibility to the RNase activity depends on several factors including intrinsic RNA properties (*e.g.* RNA secondary structures), the rate of translation, and as already discussed, RNA binding proteins and sRNAs[91,92]. First, RNA molecules arising from transcription harbor a 5' triphosphorylated (5' PPP) end, which exerts a protective role against degradation as restrains the activity of several endo or exoRNases. The 5' PPP can be converted to 5' monophosphorylated (5' P) by a pyrophosphohydrolase enzyme (RppH) that eventually triggers transcript degradation[93,94]. Second, newly synthesized transcripts can be modified at the 5' end with the addition of a nicotinamide-adenine dinucleotide cap, which prevents transcript degradation[95,96]. RNA structures at the 5' end of transcripts have been shown to impede the targeting by RNases[97,98]. At the 3' ends, the transcripts are also protected from exoRNase degradation by transcriptional terminator structures[99]. In addition, RNase target recognition and processing are often inhibited by ribosomes, which mask the RNase cleavage site on the RNA molecules and therefore promote RNA stabilization[100–102]. Lastly, RNA cellular localization also plays a role in determining the rate of degradation, as was described in *E. coli*. In this bacterium, the mRNAs encoding inner-membrane proteins are enriched at the membrane and they display a higher turnover compared to the mRNAs encoding cytoplasmic and outer-membrane proteins not localized at the membrane[103].

RNases can be classified based on their capability to cleave an RNA molecule internally (endoRNases) or from the extremities (exoRNases). EndoRNases can cleave specifically ssRNA or dsRNAs. exoRNases are further classified based on the directionality of the RNA degradation, which can occur starting from the RNA 5' end (5'-to-3' exoRNases) or from the RNA 3' end (3'-to-5' exoRNases). When the cleavage of the RNA backbone occurs through a H_2O-mediated nucleophilic attack and causes the release of a nucleoside monophosphate, the RNase is defined as hydrolytic. Conversely, RNases that use inorganic phosphate to cleave the phosphodiester bond and liberate nucleoside diphosphate are classified as phosphorolytic enzymes[104].

RNA degradation

RNA degradation (or RNA decay) allows bacteria to rapidly and constantly adjust the levels of RNAs in order to produce the correct amount of proteins needed, adapt to changing environmental conditions and recycle nucleotides. Bacteria displays a rapid mRNA turnover with a half-life of approximately 6.8 minutes in *E. coli* [15] or less than 7, 5 and 2 minutes in *B. subtilis* [98], *S. aureus* [105] and *S. pyogenes* [106], respectively. Conversely, stable RNAs including rRNAs and tRNAs, which represent the majority of the intracellular RNAs, are usually not degraded at standard growth conditions; while they undergo degradation during certain environmental stresses (*e.g.* starvation and stationary growth phase)[107]. RNA degradation mechanisms have been mainly investigated in the two model organisms *E. coli* and *B. subtilis*, which harbor different sets of RNases[108,109]. Two pathways, which differ on the initiation step, have been described and named "5' end-dependent pathway" and "direct entry pathway" (Figure 1).

5' end-dependent pathway. In both *E. coli* and *B. subtilis,* the 5' end-dependent pathway relies on the enzymatic conversion of the transcript 5' PPP to 5' P, which renders the RNA molecule susceptible to RNA degradation[93,110–112]. In *E. coli*, RNase E, which is the major endoRNase in Gram-negative bacteria, preferentially targets 5' P transcript and, after sensing the 5' P group at the RNA terminus, catalyzes multiple processing events in the targeted RNA[113,114] (Figure 1). Further, cleavage sites within the ssRNA regions are found by linearly scanning the target from the 5' P to the 3' end[115]. RNase E is an essential enzyme and plays a crucial role in RNA degradation, but does not have a homolog in Gram-positive bacteria, with the exception of some species such as *Streptomyces coelicolor* (class of Actinomycetales), *Bacillus halodurans* (class of Bacilli) and *Clostridium acetobutylicum* (class of Clostridia)[84,116]. Therefore, other RNases are involved in the initiation of RNA degradation, including the 5'-to-3' exoRNase J1 that is able to degrade transcripts with a 5'-to-3' directionality[84,117]. Upon removal of the 5' PPP group, RNase J1 recognizes the 5' P group and initiates the degradation of the transcript[94,118]. Alternatively, RNase Y, which also prefers 5' P RNAs[119], can internally cleave the target RNA and initiate degradation[108].

Direct entry pathway. RNA degradation can also be initiated directly through an internal cleavage by RNase E and RNase Y in *E. coli* and *B. subtilis*, respectively[108,120]. In this case, the activity of the endoRNases is independent from the chemical nature of the transcript 5' end (Figure 1). Indeed, it has been shown that the recognition of the 5' P is not required for the processing of several RNase E targets[93,121]. This RNA decay pathway has also been identified in Gram-positive bacteria, where RNase Y cleaves the body of the target RNAs[102,122].

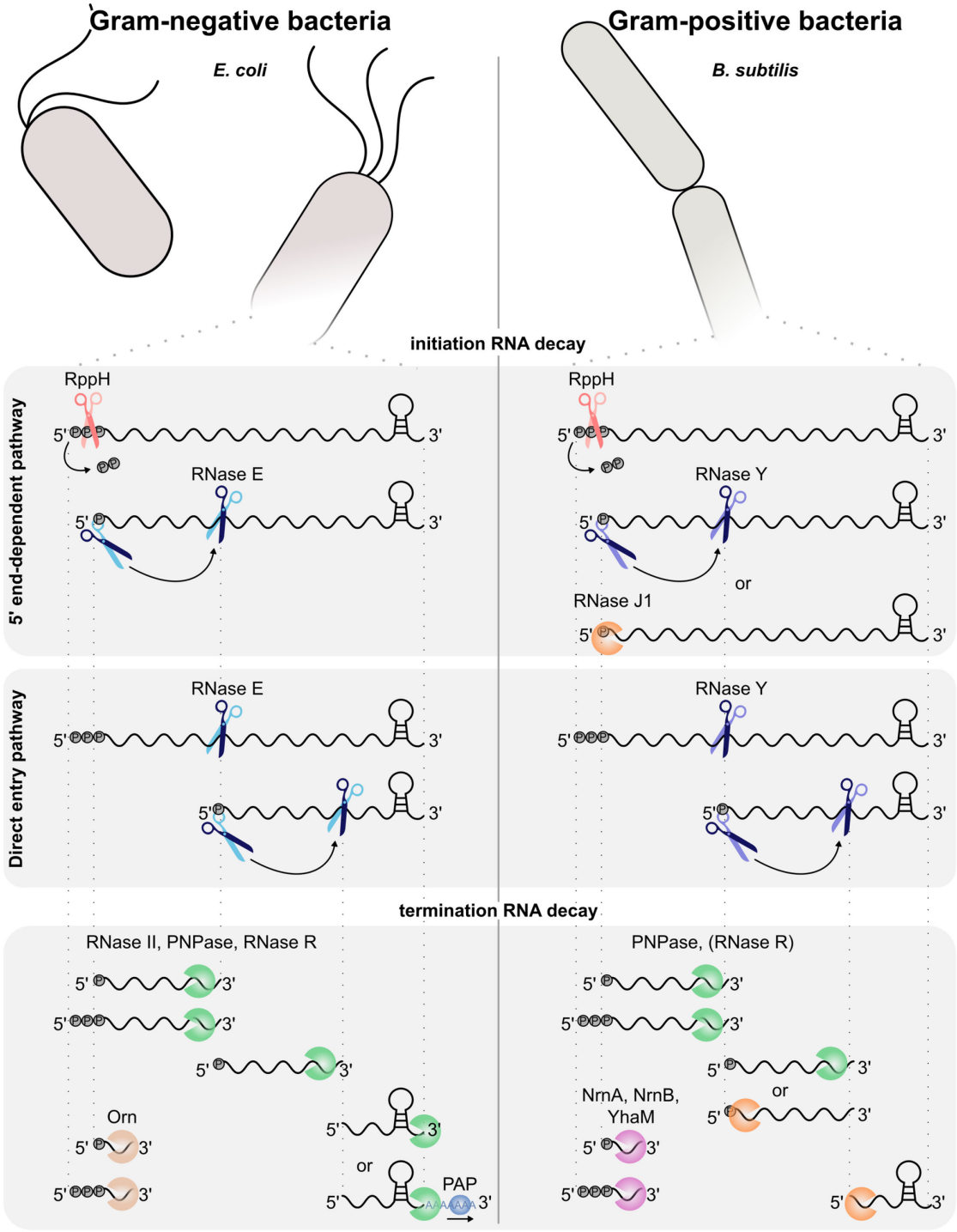


Figure 1. **RNA degradation pathways in Gram-positive and Gram-negative bacteria.**
Schematic representation of the *E. coli* (left) and *B. subtilis* (right) mechanisms of RNA degradation. In the 5' end-dependent pathway, the 5' triphosphate group (5' PPP) is converted to 5' monophosphate (5' P) by a pyrophosphohydrolase (RppH, pink scissors). In *E. coli*, this event triggers the activity of endoRNase E (blue scissors) that catalyzes multiple processing events. In *B. subtilis* the 5' P RNA is targeted either by endoRNase Y (purple scissors) or by RNase J1 (orange "pacman" symbol), which degrades the RNA from the RNA 5' to the 3' end. In the direct entry pathway, RNase E and RNase Y independently from the phosphorylation state of the transcript 5' end, cleave internally the RNA target. In *E. coli*, the degradation of the decay intermediate fragments generated by endoRNases are then degraded by 3'-to-5' exoRNases RNase II, PNPase and RNase R (green "pacman" symbol), which generate nanoRNAs further degraded by

the oligoRNase Orn (brown "pacman" symbol). In some cases, a poly(A) polymerase (PAP) (blue circle) catalyzes the addition of a poly(A) tail at the RNA 3' end, which is followed by 3'-to-5' exoRNase degradation. In *B. subtilis*, 3'-to-5' exoRNases (mainly PNPase) and RNase J1 degrade the decay intermediate fragments. The nanoRNAs are degraded by NrnA, NrnB or YhaM (pink "pacman" symbol).

RNA degradation termination. Initiation of RNA degradation by endoRNases leads to the production of decay intermediate fragments that are consequently and rapidly degraded by exoRNases[120,123] (Figure 1). In *E. coli*, the 3'-to-5' exoRNases, mainly RNase II and PNPase, are involved in degradation of the decay intermediate fragments generated by endoRNases[124–126]. These exoRNases do not complete the degradation of the fragment up to the fragment 5' end, but they produce 2-5 nt-long oligoribonucleotides (named nanoRNAs)[127]. These short RNA products are then subsequently degraded by an essential 3'-to-5' oligoRNase (Orn), which generates mononucleotides[128,129] (Figure 1). The RNA fragments containing a stem-loop terminator structure at the 3' end are usually degraded by RNase R, which displays an intrinsic helicase activity[130] or by PNPase acting in concert with RNA helicases[131,132]. Alternatively, the repetitive addition of a poly(A) tail at the RNA 3' end, by a poly(A) polymerase (PAP), renders the fragments more susceptible to 3'-to-5' exoRNase degradation[133,134] (Figure 1). Indeed, these enzymes require a ssRNA overhang at the 3' end of the substrate to bind the RNAs and perform degradation[135,136].

B. subtilis does not have an RNase II ortholog and RNA degradation is mainly performed by PNPase and to a lesser extent by RNase R[91,104,123]. *B. subtilis*, as opposed to *E. coli* harbors a 5'-to-3' exoRNase (RNase J1), which is involved in degradation of decay intermediate fragments. In particular, fragments containing the transcription terminator structure were shown to be efficiently degraded by this 5'-to-3' exoRNase[137]. RNAs with a poly(A) tail at the 3' end were also identified in *B. subtilis*, however the impact of the poly(A) on the RNA stability is unclear[138]. Finally, two nanoRNases (nanoRNase A and nanoRNase B) are able to degrade the nanoRNAs produced by exoRNases to mononucleotides. In addition, the 3'-to-5' exoRNase YhaM also possibly plays a role as nanoRNase in *B. subtilis*[139].

RNA maturation

RNases were considered to act only as degradative enzymes, but it was later found that they play a crucial role also in RNA maturation *i.e.* processing event that results in the production of a functional RNA molecule. Some RNAs, are at first produced as precursor transcript that undergoes to RNase processing responsible for the generation of a functional RNA molecule. For example, rRNAs (16S, 23S and 5S) and some tRNAs are cotranscribed and multiple endoribonucleolytic processing and trimming events are required to generate the individual rRNA molecules[140,141]. For instance, in both *E. coli* and *B. subtilis*, the 16S rRNA is released from the precursor transcript by

RNase III processing[142–144]. The mature tRNA 5' end is generated in both model organisms by RNase P[145,146], which is a ribonucleoprotein complex composed of an RNA molecule exerting the catalytic activity and a protein subunit[147]. The RNA component of RNase P also requires a maturation event in order to be active, and this is achieved by RNase E[148] and RNase Y[149] processing in *E. coli* and *B. subtilis*, respectively.

In bacteria, the production of the correct amount of proteins encoded by genes expressed from an operon, is often achieved by endoRNase-mediated processing within the polycistronic transcript. This maturation event allows differential stability of the processed mRNAs and thus to uncouple the expression of the different genes from the same operon[150–153].

In some cases, sRNAs also require a maturation step in order to be functionally active. For instance, the mature 6S sRNA, able to modulate the RNAP activity, is generated from a longer RNA precursor processed by RNase E[154,155]. This endoRNase was also shown to be highly involved in the generation of functionally active sRNAs from processing of mRNA 3' UTRs[29,30]. RNase Y also takes part in sRNA maturation, for example the RNase Y processing of the sRNA RoxS results in a truncated sRNA variant, which interacts with an increased number of target mRNAs than the precursor sRNA[156]. Finally, as already mention in section 1.1.1, the activation of CRISPR RNAs from the CRISPR-Cas systems relies on a processing event, which is performed by RNase III[53].

1.1.4 Ribonucleases involved in RNA metabolism in Gram-positive bacteria

Gram-positive bacteria harbor a different set of RNases compared to Gram-negative, with RNase Y being the major endoRNase initiating RNA degradation in *B. subtilis* and most of Firmicutes[84,109]. Although RNases are quite conserved in different Gram-positive bacteria, they often display different characteristics and functions among the various bacterial species. The main RNases involved in RNA metabolism and the regulation of gene expression in Gram-positive bacteria are discussed in detail below.

RNase Y

General characteristics. RNase Y is a ssRNA specific endoRNase, which harbors different domains: (i) a trans-membrane domain at the N-terminal region responsible for the RNase Y localization at the membrane (similar to what was described for RNase E), (ii) a disordered domain overlapping with a predicted coiled-coil region able to interact with other proteins, (iii) a ribonucleoprotein K homology (KH) domain involved in RNA binding, (iv) a histidine and aspartate (HD) containing domain responsible for the catalytic activity and (v) a C-terminal domain whose function is currently unknown[157]. RNase Y is considered to act as the

RNase E functional equivalent in the majority of Gram-positive bacteria, in which RNase E is not present. Interestingly, RNase Y can be found in *Helicobacter pylori*, a Gram-negative bacterium, and in the class of Clostridia, in which also RNase E is present[84,158].

RNase Y proteins form a dimeric structure[157,159]. RNase Y is also able to interact with other proteins involved in RNA decay and the role of RNase Y as a central enzyme of a degradosome-like complex in Gram-positive bacteria will be discussed in section 1.1.5.

Deletion of the RNase Y trans-membrane domain is lethal in *B. subtilis*[157] and it has a severe effect on *S. aureus* growth[160]. Surprisingly, in *S. aureus* the number of RNase Y targets did not vary when the trans-membrane domain was removed[160], similarly to what was observed in *S. pyogenes*[161]. Therefore, the physiological function for RNase Y membrane localization is unclear.

Effects of gene deletion. Depletion of *rny* (encoding RNase Y) expression has severe consequences on *B. subtilis* growth, cell morphology and sporulation[162,163]. While the three independent transcriptomic studies performed in *B. subtilis* revealed that depletion of RNase Y affects the expression of 20% to 30% of the transcriptome, they observed few common differentially expressed genes[157,164,165]. In *Clostridium perfringens,* the RNase Y knockout strain (Δ*rny*) could not be generated and the depletion of the enzyme severely affects the growth and the expression of up to 400 genes[166]. Oppositely, deletion of RNase Y in *S. aureus* does not have major effects on the bacterial growth and transcript abundance[167]. In *S. pyogenes*, two transcriptomic studies demonstrated that RNase Y participates in RNA degradation. In the first study, *rny* deletion caused only a small delay in growth and affected the expression of approximately 30% of the transcripts at stationary phase of growth[168]. The second showed that 98% of transcripts were stabilized in the Δ*rny* strain[106]. As RNase Y is highly conserved protein, the reason why RNase Y exerts such a different impact in these Gram-positive bacteria remains unclear. The main sequence variability among RNase Y orthologs is found in the N-terminal region of the protein, between the transmembrane domain and the disordered domain, and it might be responsible for the different RNase Y ortholog activity and specificity[84].

Progressing determinants. The study of RNase Y activity on the S-adenosylmethionine riboswitch revealed that this endoRNase preferably cleaves 5' P RNA in A/U rich regions[119], as was reported for RNase E. Only two RNA sequencing-based studies were performed to annotate RNase Y processing sites in *S. aureus* and *B. subtilis* and in both approximately one hundred cleavage sites were mapped[153,160]. RNase Y was shown to preferentially cleave after a guanosine (G) residue in A/U ssRNA regions in *S. aureus*[160]. The role of this nucleotide is unclear, as it is not required for the *saePQRS* transcript processing by RNase Y, which instead depends on the presence of a stem-loop structure located 6 nt downstream of the RNase Y processing site[169]. In *B. subtilis*, the mapping

of the RNA 5' ends generated by RNase Y did not reveal any sequence or structural preference[153]. In this study, RNase Y was shown to form a complex with three other proteins (YlbF, YmcA and YaaT), which affect RNase Y specificity[153]. This complex was previously shown to be involved in the regulation of biofilm formation, competence and sporulation in *B. subtilis*[170–173]. The complex mediates the destabilization of the transcript encoding the master regulator of biofilm formation (SinR), by promoting RNase Y cleavage[173,153]. Recently, it has been shown that the YlbF, YmcA and YaaT complex consists of an heterotrimer harboring a double iron-sulfur cluster, which is ligated by cysteines[174]. These residues were demonstrated to be necessary for the maturation processing of the *cggR-gapA* operon by RNase Y[174].

Functions. RNase Y is involved in uncoupling the expression of genes that are encoded within the same polycistronic transcript[159,175], in the maturation of the RNA component of RNase P[106,149], in the decay of premature terminated transcripts derived from regulatory 5' UTRs (*e.g.* riboswitches and T-boxes)[119,153,160] and in modulating the abundance of regulatory sRNAs[106,160,166,167]. Interestingly, RNase Y is required for the virulence of *S. aureus*[176,177], *S. pyogenes*[168] and *C. perfringens*[166]. For instance, RNase Y is responsible for stabilizing a transcript encoding the two-component system SaeRS, which regulates the expression of numerous virulence genes in *S. aureus*[167,169]. Attenuated virulence was observed for the Δ*rny* strain in both silkworm and murine *S. aureus* infection models[176]. In *C. perfringens*, RNase Y controls the expression of the κ-toxin collagenase by acting in concert with a sRNA[166]. In *S. pyogenes*, RNase Y affects the expression of virulence genes during stationary phase of growth in accordance to the nutrient availability (*e.g.* carbon and peptides)[168] and it is required for the thermoregulation of capsule production[178]. An impaired subcutaneous infection was observed after injection of the Δ*rny* strain in mice, indicating that RNase Y is required for *S. pyogenes* pathogenesis[168].

RNase III

General characteristics. RNase III is a double-stranded (dsRNA) specific endoribonuclease, which cleaves the phosphodiester bond between two nucleotides leaving a 3' hydroxyl (OH) and 5' P termini and 2 nucleotide (nt) 3' overhang[179,180]. RNase III functions as a homodimer and the catalytic site of each subunit is responsible for cleaving independently one of the RNA strands, generating a double strand break[179,180]. However, RNase III can also nick RNA, meaning cleaving only one RNA strand of the duplex[180–182]. The gene encoding for the bacterial RNase III (*rnc*) is not essential, with the exception of *rnc* from *B. subtilis*. In this bacterium, RNase III is required to repress the production of toxins encoded in prophage regions, as it cleaves

the RNA duplex formed by the toxin mRNAs and its respective RNA antitoxin[164]. An RNase III deletion strain could be constructed only after the deletion of these prophage regions[84,164].

Functions and effects of gene deletion. RNase III plays a crucial role in rRNA maturation by producing the precursor transcript of the 16S and 23S rRNAs[141]. It is also involved in the regulation of gene expression, often in concert with regulatory sRNAs[84,92]. Deletion of *rnc* in *B. subtilis* affects the expression of 12% of the transcripts, similar to what was previously observed in *E. coli*[164,183]. Mapping of the RNase III cleavage sites by RNA sequencing[137] and co-immunoprecipitation of RNase III followed by RNA sequencing[184] revealed that RNase III, apart from rRNA, targets also mRNAs and several intergenic regions in *B. subtilis* and *S. aureus*, respectively. In addition, in *S. aureus* the amount of antisense transcription was increased in the Δ*rnc* strain compared to the WT strain, indicating an important role for RNase III in antisense RNA regulation[16,184]. Dissimilarly, in *B. subtilis* and *S. pyogenes* the effect of RNase III on antisense RNA processing and clearance seemed to be minor[137,185].

In *S. pyogenes*, RNase III is involved in the activation of the crRNAs from the type II CRISPR-Cas system (see section 1.1.1). The tracrRNA harbors a complementary region to the pre-crRNA and the duplex formed by the two RNAs is recognized and cleaved by RNase III. This processing event results in the maturation of the crRNA, which in complex with tracrRNA, guides the nuclease Cas9 to the invading nucleic acids[53,54]. Genome wide mapping of the RNase III cleavage sites in *S. pyogenes* revealed that this enzyme preferentially cleaves in UTRs rather than in CDSs[185], as was also demonstrated in *S. aureus*[184]. For instance, the processing of RNase III in the UTR between *secY* (encoding Sec translocase) and *adk* (encoding adenylate kinase) uncouples the expression of the two genes[185]. In this bacterium RNase III displays a broad nicking activity *in vivo*, whose role in transcript regulation is not known yet. Overall, at standard growth conditions, deletion of *rnc* has little impact on gene expression and antisense transcription. Overall, RNase III is mainly involved in RNA maturation pathways (*e.g.* rRNA maturation) in *S. pyogenes*[185].

RNases J1 and J2

General characteristics. In eukaryotes, RNA degradation by exoRNases occur in both 3'-to-5' and 5'-to-3' directions. The search in *E. coli* for an exoRNase capable of degrading RNA from 5' to 3' transcript ends was not successful. In 2007 a 5'-to-3' exoRNase, named RNase J1, was instead identified in *B. subtilis* and it was shown to be involved in the 16S rRNA maturation[117]. More recently, a 5'-to-3' exoRNase was also identified in *E. coli* (RNase AM), however its activity was only validated *in vitro* on RNA oligonucleotides[186].

While Gram-positive bacteria with low guanosine and cytosine content (low-GC) display two RNase J orthologs (RNase J1 and J2), the high-GC Gram-positive bacteria encodes only for one ortholog (RNase J1). These enzymes are mostly present in Gram-positive bacteria, with the exception of *H. pylori,* a Gram-negative bacterium that harbors a homolog of RNase J1 (RNase J) is expressed[84,162].

RNase J1 operates as a processive enzyme (*i.e.* removes one nucleotide at the time), but can also act as distributive enzyme when the target is shorter than 5 nt[187]. Transcripts with a 5' P or a 5' OH are the preferred substrates of RNase J1, as the enzyme cannot bind to 5' PPP RNAs[117,188]. RNase J1 is unable to degrade the transcripts whose 5' end regions are involved in the interaction with another RNA molecule[37]. Although RNase J1 activity is impaired by RNA secondary structures, this enzyme is involved in the degradation of terminator step-loops during RNA decay[189]. RNase J1 was also shown to act as an endoRNase *in vitro*, but this activity was never confirmed *in vivo*, where RNase J1 is believed to act mainly as a 5'-to-3' exoRNase[190]. In *B. subtilis* RNase J2 exoribonucleolytic activity is much weaker when compared to the one of RNase J1 and the two enzymes have been shown to interact forming a heterodimer complex[191,192]. It was proposed that RNase J2 exert more a structural role in the complex with RNase J1[190].

Functions and effects of gene deletion. Deletion of the gene encoding RNase J1 (*rnjA*) in *B. subtilis* has a drastic effect on growth, cell physiology and alters the expression of approximately 21% of the transcripts[162]. Coherent with the reduced exoribonucleolytic activity of RNase J2 (encoded by *rnjB*), the deletion of *rnjB* did not affect cell growth in this bacterium. Different impacts of RNases J1 and J2 on the bacterial physiology and gene expression have been observed in other Gram-positive bacteria than *B. subtilis*. For instance, in *S. aureus* deletion of *rnjB* causes stronger growth defects than *rnjA* deletion[193]. In *Streptococcus mutants*, these two RNases are not essential and RNase J2 acts as an endoRNase *in vivo* independently from RNase J1[194]. It was shown that RNase J2 cleaves the *gbpC* mRNA and the processing can be inhibited by the base-pairing of *gbpC* mRNA with the 5' UTR of *irvA* mRNA. Lastly, in *S. pyogenes*, both RNases are essential and they exert a role in the regulation of transcript stability[195].

PNPase

General characteristics. PNPase (polynucleotide phosphorylase) is a phosphorolytic 3'-to-5' processive exoRNase that uses inorganic phosphate to degrade RNAs and thereby releasing nucleoside diphosphate. PNPase was also shown to catalyze the reverse reaction, meaning that it is able to synthesize RNA from nucleoside diphosphates without a template[196,197]. The major activity of PNPase *in vivo* is the degradation of RNAs more than their synthesis, but the polymerase

function is used to add polynucleotide tails at the RNA 3' ends to favor RNA degradation by 3'-to-5' exoRNases. However, this role appears to be less important in *B. subtilis* than in *E. coli*[138,198]. As already mentioned, PNPase requires a ssRNA tail of at least 6 nt at the RNA 3' ends to bind the target and perform RNA degradation[199,200]. PNPase displays a ring-like structure, which forms a channel ending with the catalytic site. The channel can accommodate only ssRNA, explaining why this enzyme is unable to digest secondary structures[109]. Therefore, PNPase often acts in concert with the CshA helicase and the two enzymes were shown to physically interact[88].

Functions and effects of gene deletion. In *B. subtilis*, PNPase functions as the major 3'-to-5' exoRNase involved in RNA degradation and deletion of the corresponding gene (*pnpA*) causes pleiotropic effects (*e.g.* increased sensitivity to antibiotics and cold, impaired competence)[123,201–203]. In *S. pyogenes*, PNPase was also demonstrated to act as the main 3'-to-5' exoRNase in RNA decay, but the Δ*pnpA* strain does not exhibit major phenotypes except for reduced fitness at low temperatures[204]. In both *B. subtilis* and *S. pyogenes*, PNPase is involved in the degradation of decay intermediate fragments generated by endoRNases[136,204].

In addition, PNPase was shown to participate in sRNA regulation in *E. coli* by sequestering the Hfq-dependent sRNAs thereby protecting them from degradation[205]. This activity has never been observed in Gram-positive bacteria. However, in *L. monocytogenes* PNPase is required for the maturation of the CRISPR RNA RliB and for the RliB activity in the defense against invading nucleic acids[206].

RNase R

General characteristics. RNase R is a hydrolytic 3'-to-5' exoRNase that degrades RNA in a processive mode and also displays an RNA helicase activity that is involved in the unfolding of secondary structures[207,208]. *In vitro*, RNase R is able to digest long dsRNA regions, but a ssRNA tail at the 3' end (optimally 10 nt-long) is necessary for the enzyme to degrade the substrates[130,130,209].

Functions and effects of gene deletion. Deletion of the gene encoding RNase R (*rnr*) in *B. subtilis* does not affect the bacterial growth and bulk RNA stability[136]. The double *rnr* and *pnpA* deletion strain grows slower than the single *pnpA* mutant, indicating that RNase R can in some cases compensate for the absence of PNPase. In *S. pyogenes*, however, *rnr* and *pnpA* could not be simultaneously deleted, indicating that this bacterium requires the presence of at least one of the two 3'-to-5' exoRNases[204].

RNase R is involved in the maturation of stable RNAs (*e.g.* rRNAs) and in RNA quality control by degrading fragments derived from defective transcripts[130,210]. RNase R also participates in RNA decay, ensuring the degradation of structured fragments[136,211]. However, the overall role of

RNase R in RNA degradation is limited, as was shown in both *B. subtilis* and *S. pyogenes*[136,204]. Instead, RNase R exerts a more critical role under non-standard growth conditions. For instance, in *E. coli* RNase R production is triggered during environmental stresses (*e.g.* cold)[212,213]. RNase R was shown to be involved in the adaptation to cold shock by degrading transcripts during the acclimation phase and allowing translation recovery[214]. Coherently with the prominent role of RNase R during stress, RNase R production is tightly controlled. The regulation of this RNase will be further discussed in section 1.1.6.

YhaM

General characteristics. YhaM is a hydrolytic 3'-to-5' exoRNase found only in Gram-positive bacteria[215]. The enzyme is composed of an oligonucleotide/ oligosaccharide binding (OB) domain and a HD domain, which is responsible for the catalytic activity[215].

Functions. This exoRNase degrades nanoRNAs *in vitro* and overexpression of the gene encoding YhaM (*yhaM*) was able to partially compensate for the absence of the oligoRNase Orn in *E. coli* [139]. It has been proposed that YhaM, by removing few nucleotides at the transcript 3' ends, could prevent degradation of the transcripts by RNase R, which requires a sufficiently long ssRNA tail to degrade the substrate[136]. YhaM was also shown *in vitro* to degrade longer RNA substrates and to be sensitive to secondary structure[215]. YhaM ortholog in *S. aureus* (CBF1), beside displaying RNA degradation activity *in vitro*[215], has the ability to bind DNA and it was shown to participate in DNA replication[216]. In *B. subtilis* the production of this enzyme is induced under conditions of DNA damage[217] and it degrades DNA in the 3'-to-5' direction *in vitro*[215], suggesting that YhaM could play a role for DNA repair mechanisms. Deletion of *yhaM* does not affect the bacterial growth and global transcript levels neither in *B. subtilis* nor in *S. pyogenes*[204,215]. Recently, it has been shown in *S. pyogenes* that YhaM acts as an unspecific 3'-to-5' exoRNase that removes on average 3 nt at the RNA 3' ends of the majority of the transcripts in the cell. The biological consequences of such general trimming activity are still unclear[204]. Interestingly, deletion of *yhaM* causes growth defect at low temperature in *S. pyogenes* and the effect is even more pronounced when both *yhaM* and *rnr* are deleted[204]. These results indicate that YhaM could play an important role under specific environmental conditions. Overall, the function of YhaM in Gram-positive RNA metabolism remains elusive.

NanoRNases

3'-to-5' exoRNases do not complete the digestion of RNAs, but usually generate short RNA products (nanoRNAs or oligoRNAs), which then are fully digested by nanoRNases (or

oligoRNases). In *E coli*, this function is performed by Orn, an essential oligoRNase that is not present in *B. subtilis*. Instead, this bacterium harbors two nanoRNases (NrnA and NrnB) and the double deletion mutant of these two genes is viable[127,139,218]. Moreover, simultaneous deletion of the genes encoding NrnA, NrnB and YhaM is not lethal indicating that in *B. subtilis* the degradation of nanoRNAs is probably performed by enzymes acting redundantly[139]. NrnA is a bidirectional exoRNase able to degrade from the 5' extremity a 12-nt long oligoRNA and to digest the substrate from the 3' extremity when the oligoRNA is not longer than 3 nt[219]. NrnB degrades 5-nt oligoRNAs *in vitro* with a 3'-to-5' directionality[139]. Some Gram-positive bacteria, like *S. aureus* and *S. pyogenes*, only encode NrnA[84].

1.1.5 RNA degradosome and localization of RNA degradation

In *E. coli* and other Gram-negative bacteria, the enzymes involved in RNA degradation are part of a multiprotein complex called degradosome. The central factor of this complex is RNase E, which is anchored at the bacterial membrane and through its C-terminal domain is able to interact with PNPase, RhlB and enolase (an enzyme involved in the glycolytic pathway)[220,221]. The degradosome components can change upon variations of the environmental conditions and stimuli. For instance, in some circumstances the Hfq-sRNA complex contacts RNase E in order to promote the sRNA-mediated degradation of the target mRNAs[158].

It was proposed that a degradosome-like complex is also formed in Gram-positive bacteria, with RNase Y replacing RNase E. However, the existence of this complex is still controversial[222,223]. In *B. subtilis*, RNase Y by two-hybrid assays was shown to interact with enolase and phosphofructokinase, RNA helicase CshA, RNase J1 and PNPase[159]. The interaction of RNase Y with enolase and CshA was also confirmed using different methods in both *B. subtilis* and *S. aureus*[88,157,224]. However, the interaction of RNase J1 and RNase Y was not observed by co-purification experiments[192,225] and RNase J1 was instead shown to interact with RNase J2 and CshA, which are both associated to the ribosomes[191,225]. In *S. aureus*, RNase J2 was recently shown to interact not only with RNase J1, but also with several other proteins including RNase Y, CshA and enolase. In addition, it also transiently forms complexes with DnaK, DnaJ (two proteins involved in protein quality control) and FtsZ (a GTPase involved in cell division)[226].

The interaction of RNase Y with PNPase, although not confirmed in *S. aureus*, likely occurs in *B. subtilis*. The residues involved in this interaction were identified and resembled the ones used by PNPase to interact with RNase E in *E. coli*[227]. However, mutations of these residues do not affect the expression of two known RNase Y-PNPase targets in *B. subtilis*[227].

The only protein interaction of RNase Y reported in *S. pyogenes* is the one with enolase[168], which was also identified in *S. aureus* and *B. subtilis*[159,224]. The reason why enzymes involved in the carbon metabolism are associated with the RNA decay machinery remains unclear. It has been proposed that these enzymes could play a structural role in the degradosome or could regulate the degradation of RNA molecules based on the metabolic state of the cell[222,228].

More recently, a localization study of the putative degradosome complex components revealed that while RNase Y is localized at the membrane, the putative interacting partners (*i.e.* PNPase, enolase and phosphofructokinase) were found uniformly distributed in the cytoplasm[229]. It is possible that the interactions between the components of the degradosome-like complex are transient and do not occur simultaneously, as the entire complex has never been purified in Gram-positive bacteria[221–223].

The reason why RNase E and RNase Y are anchored at the membrane, with the consequence of compartmentalizing the RNA degradation at the cell periphery, is unclear. In *E. coli* it has been recently shown that the RNase E membrane localization is required for RNase E protein stabilization, assembling of the degradosome and protecting the newly synthesized transcripts from premature degradation[230].

1.1.6 Post-transcriptional regulation mediated by RNases

RNases can regulate gene expression at the post-transcriptional level by targeting specifically an mRNA, leading to its degradation or stabilization. The targeting of a specific mRNA can be achieved by modifying the RNase accessibility to the mRNA, for instance trough the action of sRNAs. Alternatively, the activity of RNases can also be modulated by interacting proteins that promote the degradation or stabilization of certain mRNAs, in response to external stimuli. Examples of RNase activity regulation are described below in detail.

RNases and sRNAs

A plethora of sRNA-RNase mechanisms of gene expression regulation have been described in Gram-positive bacteria[92,231] (Figure 2). sRNAs can modulate RNase activity in order to promote either target mRNA degradation (Figure 2A and B) or stabilization (Figure 2C and D). The two outcomes can be a consequence of either an indirect sRNA effect on the mRNA stability through the regulation of translation initiation (Figure 2A and C), or a direct role on mRNA stability by affecting RNase processing (Figure 2B and D). In the following paragraphs examples of sRNA and RNase concerted action in gene expression regulation are discussed.

sRNA-mediated RNA degradation by translation inhibition. *In B. subtilis*, the sRNA RoxS, which is involved in the response to oxidative stress, was shown to bind to the RBS of the *ppnkB* mRNA, encoding a nicotinamide adenine dinucleotide kinase[156]. This interaction negatively affects the mRNA translation as it blocks the access of the ribosome to the *ppnkB* mRNA. The reduced ribosome trafficking on the mRNA facilitates the processing by RNase Y[156] (Figure 2A). In addition, the dsRNA formed by RoxS and *ppnKB* mRNA is cleaved by RNase III. The overall output of these two processing events is the rapid *ppnKB* mRNA degradation (Figure 2A).

sRNA-mediated RNA degradation by RNase processing. The *B. subtilis* genome encodes several type I TA systems consisting of an mRNA (toxin) and asRNA (antitoxin) produced from the same locus. In the *txpA/ratA* system, the asRNA RatA binds the *txpA* toxin mRNA and triggers RNase III processing of the duplex region[164] (Figure 2B). The products deriving from the RNase III cleavage are subsequently degraded by PNPase and RNase J1 leading to downregulation of the toxin expression[164] (Figure 2B). The processing of RNase III is essential to ensure bacterial viability, this is the reason why *rnc* can be deleted in *B. subtilis* only in the absence of type I TA systems regulated by RNase III.

sRNA-mediated RNA degradation by translation activation. The sRNA VR-RNA from *C. perfringens*, binds at the 5' end region of the *colA* mRNA, coding for a collagenase involved in virulence[166,232]. Prior VR-RNA binding, the 5' UTR of *colA* mRNA harbors a stem loop structure, which sequesters the RBS and renders the mRNA unstable (Figure 2C). The binding of VR-RNA induces a change in the secondary structure of the *colA* mRNA 5' UTR, thereby triggering RNase Y cleavage and subsequent liberation of the RBS (Figure 2C)[166]. The processing of RNase Y is responsible for another structural change in the *colA* mRNA 5' UTR. A stem-loop structure is formed after RNase Y processing of the *colA* mRNA 5' UTR and it protects the *colA* mRNA from degradation (Figure 2C).

sRNA-mediated RNA stabilization by inhibition of RNase processing. In *S. pyogenes*, the FasX sRNA binds and forms a duplex with the 5' extremity region of *ska* mRNA (Figure 2D), coding for a streptokinase involved in *S. pyogenes* pathogenesis[233]. This interaction likely prevents the activity of RppH and/or RNase J1, whose activities are both inhibited by dsRNA at the 5' end region of the transcripts[233] (Figure 2D). Therefore, in the presence of FasX the stability of *ska* mRNA is increased because its degradation is blocked. It has been proposed that in the absence of FasX, the 5' PPP of the *ska* mRNA is targeted by RppH, which generates a 5' P group recognized by RNase J1. This model has not been experimentally confirmed, but it is known that both RNase Y and PNPase are not involved in the FasX-mediated regulation of *ska* mRNA.

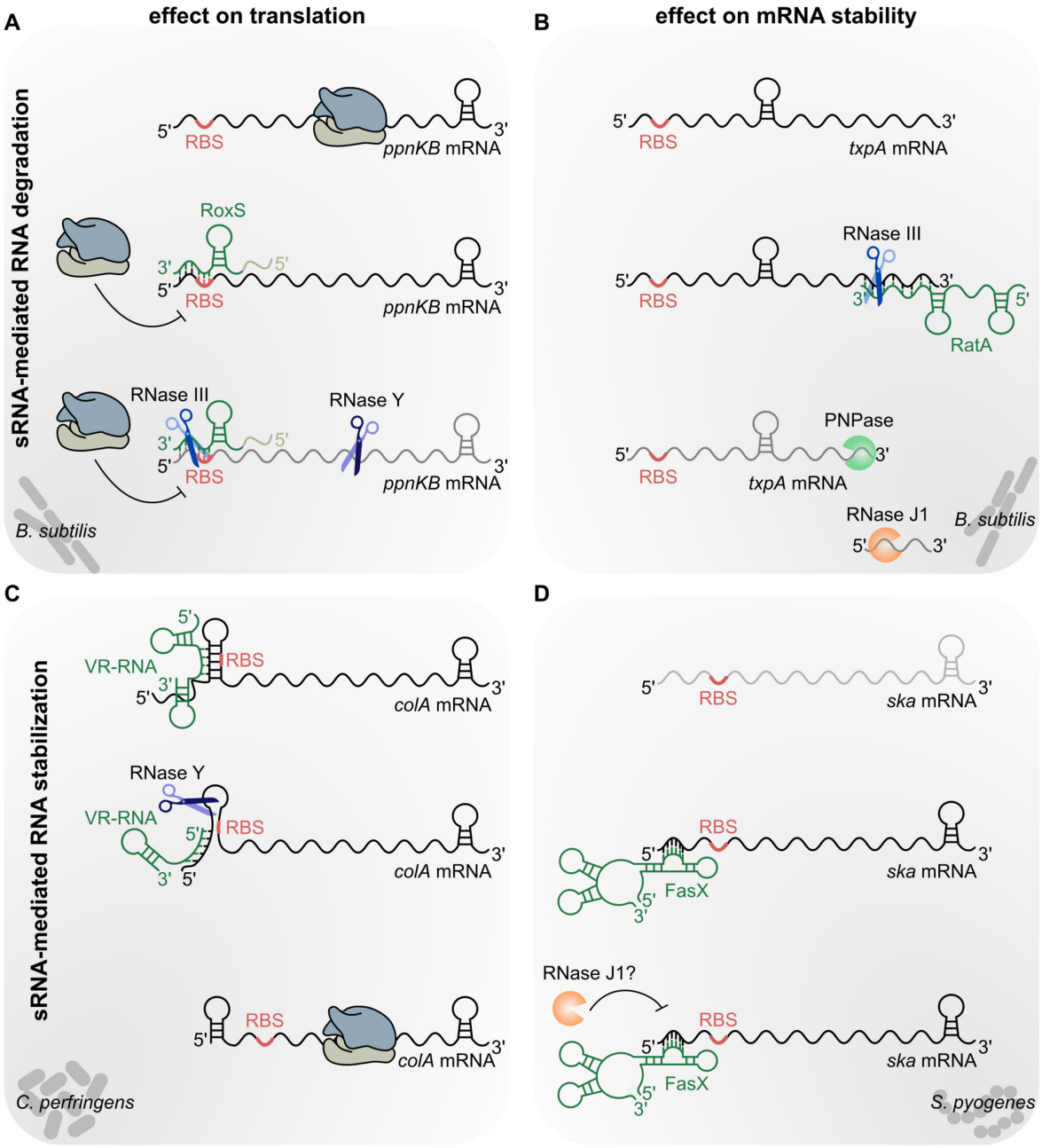


Figure 2. **Modulation of RNase activity by sRNAs.**
A. The sRNA RoxS from *B. subtilis* binds to the ribosome binding site (RBS) of the *ppnKB* mRNA and inhibits its translation. The dsRNA region formed by the interaction between *ppnKB* mRNA and RoxS sRNA is targeted by RNase III (blue scissors). In the absence of ribosomes on the *ppnKB* mRNA, the RNase Y (purple scissors) processing site is exposed and RNase Y can cleave the mRNA. **B.** In *B. subtilis*, the 3' region of the antitoxin sRNA RatA binds the *txpA* mRNA, encoding a toxin. RNase III cleaves the dsRNA leading to *txpA* mRNA degradation, which is performed by PNPase (green "pacman" symbol) and RNase J1 (orange "pacman" symbol). **C.** The VR-RNA base-pairs with the 5' end region of the *colA* mRNA in *C. perfringens*. The binding induces a conformational change that triggers RNase Y cleavage. The processing leads to an increase of the mRNA stability, because of the formation of a stem-loop structure at the transcript 5' end. **D.** The sRNA FasX from *S. pyogenes* binds the 5' end region of the *ska* mRNA blocking the mRNA degradation, likely performed by RNase J1, and leads to mRNA stabilization.

Regulation of RNase production and activity

The control of RNase expression and activity have been elucidated only for a limited number of RNases and only in few organisms[234]. Below the best characterized mechanisms of RNase expression and activity regulation are described.

RNase III. In several bacteria, RNase III negatively autoregulates its own expression, by cleaving the *rnc* mRNA leading to its degradation[137,184,235]. At post-translational level, RNase III can interact with different proteins that modulate its activity. For instance, in *E coli* during T7 bacteriophage infection, the T7 protein kinase stimulates RNase III activity to promote the processing and maturation of the phage mRNAs[236]. In *E coli*, RNase III interacts with the YmdB stress response regulator, which inhibits the RNase III dimer formation thereby reducing RNase III activity under cold-shock or at stationary growth phase[237].

RNase Y. In *B. subtilis*, RNase Y expression is growth-phase dependent[157] and in *S. pyogenes* RNase Y activity is affected by the cell nutritional conditions through an unknow mechanism[168]. In *S. aureus*, RNase Y cleaves the *rny* mRNA, likely regulating its own expression[160]. In addition, in this bacterium RNase Y interacts with flotillin (FloA), which consists of a protein-associated membrane involved in the formation of microdomains within membranes. In the absence of FloA, RNase Y dimerization and thereby activity is impaired[238].

RNases J1 and J2. In *B. subtilis*, RNase J1 appears to be able to autoregulate its own production and RNase J2 likely negatively affects the expression of RNase J1[239].

PNPase. PNPase expression and regulation are tightly regulated. PNPase can autoregulate its own expression by degrading the *pnpA* transcript after RNase III processing of the *pnpA* mRNA 5' UTR or by inhibiting *pnpA* mRNA translation[240,241]. In *E coli*, *pnpA* expression was also shown to be regulated by sRNAs. The sRNA CsrA prevents RNase III processing and therefore positively regulates *pnpA* expression[242]. In *E. coli* PNPase production is also regulated by the asRNA SraG, which induces *pnpA* mRNA destabilization likely by inhibiting translation[243]. In addition, stress conditions (*e.g.* cold-shock)[244,245] or metabolites (*e.g.* citrate, ppGpp and ATP) can affect *pnpA* expression[246–248].

RNase R. RNase R, similarly to PNPase, is tightly regulated and different stress conditions can affect the levels of this enzyme in the cell. RNase R is a highly unstable protein, but during cold shock or starvation the protein stability increases[212,213]. RNase R is post-translationally modified by the addition of an acetyl group, which renders the protein unstable during exponential growth phase[249,250]. At stationary phase of growth, RNase R is not acetylated and is no longer subject to proteolysis[249,250]. Binding of RNase R to the ribosomes also affects the levels of RNase

R; during exponential phase of growth RNase R is mainly associated to the ribosomes and the unbound RNase R proteins are degraded[251].

1.2 Methodologies to study RNases

1.2.1 Characterization of RNase targets

The investigation of the RNase properties and functions relies on the identification and characterization of their targets. RNases have been largely studied *in vitro*, using biochemical approaches on a limited number of substrates. These studies are highly valuable, but they come with several disadvantages: (i) difficulties in purifying the enzymes (especially membrane associated RNases such as RNase Y), (ii) limited number of targets tested, (iii) absence of RNase interacting proteins that could play a role in regulating the RNase activity and specificity, (iv) difference in the RNA structures formed *in vitro* and *in vivo* and (v) absence of post-translational modifications that could affect the RNase activity. Other techniques, based on the comparison of the RNA *in vivo* in the WT strain and in the deletion mutant of the RNase of interest (Δ*rnase*) (*e.g.* Northern blot analysis and primer extension), although very informative are not suitable for high-throughput studies. More recently, transcriptomic studies based on RNA sequencing gave global insights into the genome-wide transcript abundance and improved the understanding of the bacterial RNase function. RNA sequencing is a high-throughput technique which allows the simultaneous identification of the RNA sequence and abundance from a specific sample. The total RNA or a specific fraction of RNAs is converted into a library of cDNA fragments that harbor adapter sequences at one or both the RNA extremities. Each cDNA is then sequenced from one extremity (single-end sequencing) of from both the extremities (pair-end sequencing). The sequenced cDNAs (*i.e.* reads) are then mapped to reference a genome.

The steady state levels of total RNAs in the WT and Δ*rnase* strains are usually calculated to identify the transcripts that are differentially expressed in the absence of the RNase of interest[16,123,252–255]. A different approach is to inhibit transcription by blocking the RNA polymerase (using *e.g.* rifampicin). The RNA abundance is measured by RNA sequencing at different time points after the rifampicin treatment. This strategy allowed an estimation of the RNA half-life in different bacteria[106,256,257]. The limit of these approaches is the impossibility of distinguishing between direct targets (*i.e.* regulated by the RNase at post-transcriptional level) or indirect targets (*i.e.* regulated by other factors which are controlled by the RNase under study).

1.2.2 RNA sequencing methodologies to study bacterial RNase direct targets

Identification of the RNAs associated to an RNase

Several tailored RNA sequencing based methodologies have been developed in order to specifically identify the direct targets of an RNase. Pull-down of the RNAs associated to catalytically inactive RNase III followed by RNA sequencing was used to globally map the RNase III direct targets in *S. aureus*[184]. This technique exploits epitope-tagged inactive variants of RNase III yet able to bind to its substrates (Figure 3A). The target RNAs are captured *in vivo*, as they are bound to RNase III mutants unable to cleave. In the WT scenario, the fragments generated by RNase III processing will be displaced and degraded, or occasionally one of the generated products can be found associated with RNase III. In these cases, the 5' termini of the RNA pulled-down with the WT RNase III variant corresponded to the RNase III cleavage sites mapped *in vitro*[184]. RNA targets bound to the inactive RNase III variants are not necessarily direct targets of RNase III, as this enzyme could bind its substrates without cleaving[182]. For this reason, RNase III processing of the co-immunoprecipitated RNAs had to be confirmed by mapping the cleavage sites *in vitro*[184].

A similar strategy was applied to identify sRNA-mRNA duplexes associated with RNase E in *E. coli* and the method was named "crosslinking ligation and sequencing of hybrids" (CLASH)[258] (Figure 3B). In the CLASH approach RNAs are cross-linked to a flagged active RNase E, by *in vivo* treatment with ultraviolet radiations (UV). Subsequently, the covalently bound RNAs are trimmed and RNA linkers are added before the cDNA library preparation[258] (Figure 3B). This approach allows not only to identify at the single nucleotide resolution the RNA region bound to an RNase, but also the RNA-RNA interactions taking place on the enzyme.

RNase III direct targets were identified in *E. coli* using a different RNA sequencing-based methodology, which relies on the specificity of the J2 antibody able to bind dsRNAs independently from their sequence[259] (Figure 3C). The RNAs immunoprecipitated with the J2 antibody from the *E. coli* WT and Δ*rnc* strains were used for the cDNA library preparation (Figure 3C). RNA duplexes were mainly detected in the Δ*rnc* strain than in the WT strain, suggesting that in *E. coli* RNase III plays a critical role in the degradation of dsRNAs[259].

Identification of the RNase processing positions genome-wide

The above-mentioned studies are not suitable to determine the exact location of the processing events. Therefore, other RNA sequencing approaches based on the genome-wide mapping of the RNA ends in the WT and Δ*rnase* strains were employed (Figure 3D). The majority of these methods rely uniquely on the mapping of RNA 5' ends and they were used to identify

processing sites of several endoRNases: (i) RNase III cleavage sites in both *B. subtilis*[137] and *E. coli*[260,261], (ii) RNase Y cleavage sites in *S. aureus*[160] and *B. subtilis*[153], (iii) RNase J1 cleavage sites in *S. aureus*[193] and (iv) RNase E cleavage sites in *E. coli*[30,121].

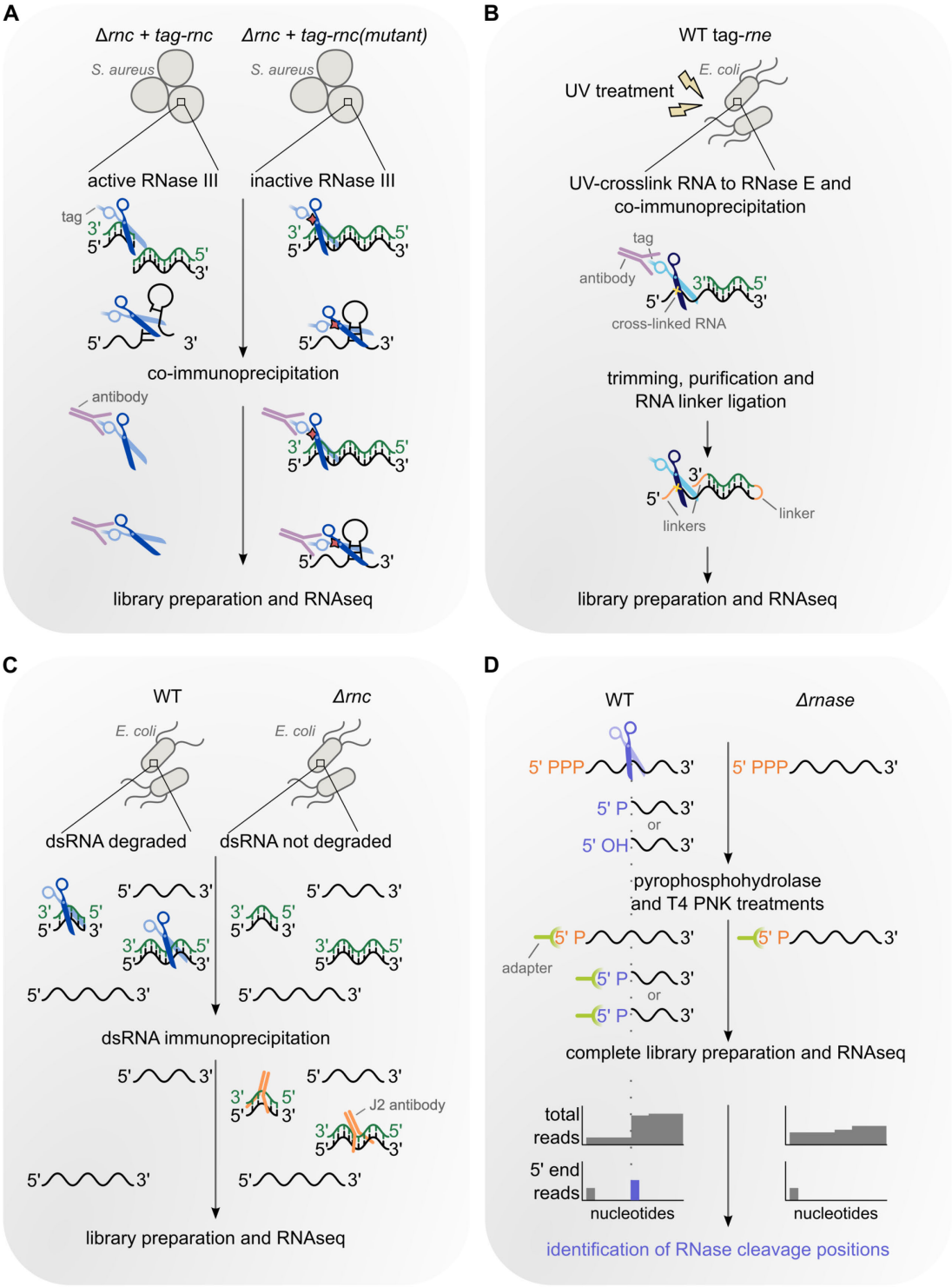


Figure 3. **Methodologies to study RNases using RNA sequencing.**
A. In *S. aureus* the gene encoding RNase III (*rnc*) was deleted and epitope-tagged WT RNase III (tag-*rnc*) or catalytically inactive RNase III (tag-*rnc* mutant) were expressed. The inactive RNase III variant (blue scissors with red star) is able to bind the RNA target but not to cleave. The RNAs bound to RNase III are co-

immunoprecipitated with the enzyme, using an antibody recognizing the epitope and used for library preparation and subsequent RNA sequencing (RNAseq)[184]. **B.** *E. coli* expressing a tagged-RNase E (tag-*rne*) was treated with UV in order to link the RNA targets (black curved line) with the interacting sRNAs (green curved line) to RNase E (blue scissors). The yellow cross indicates the covalent bound between the RNA and the RNase after UV treatment. RNase E was pulled-down with the bound RNAs using an antibody recognizing the tag and the RNAs were trimmed before ligation of the RNA linkers. These RNAs were used to prepare cDNA libraries for RNAseq[258]. **C.** In *E. coli*, RNAs from the WT and RNase III deletion mutant (Δ*rnc*) strains were isolated and incubated with the antibody J2, which specifically recognizes double-stranded RNA (dsRNA) regions. Less dsRNAs are immunoprecipitated in the WT strain compared to the Δ*rnc* strain[259]. **D.** Mapping of the RNA 5' ends in the WT strain and in the strain deleted of the RNase of interest (Δ*rnase*). An RNase (purple scissors) cleaves an RNA molecule producing a new 5' RNA end, which harbors either a monophosphate group (5' P) or a hydroxyl group (5' OH). Total RNAs from the WT and Δ*rnase* strains are treated with a pyrophosphohydrolase to covert the 5' triphosphate group (5' PPP) in 5' P and with T4 polynucleotide kinase (T4 PNK) to convert the 5' OH in 5' P. The monophosphate group at the transcript 5' ends is required during the library preparation for the adapter ligation (depicted in orange). The sequenced cDNAs, named reads, are mapped (total reads) as well as the reads starting at each nucleotide (5' end reads). The RNA 5' end detected in the WT strain, but not in the Δ*rnase* strain (purple bar) corresponds to the RNase cleavage position.

Although these studies adopt different library preparation protocols, they all depend on the ligation of an adapter sequence at the RNA 5' ends (Figure 3D). Whereas primary transcripts harbor a PPP group at the transcript 5' ends, processed transcripts harbor either a P or OH group. Since adapter ligation at the RNA 5' ends requires a 5' P, total RNAs are first treated with a pyrophosphohydrolase (*e.g.* RppH or Tobacco acid pyrophosphatase), which converts the 5' PPP into 5' P thereby enabling the sequencing of the total RNA.

To specifically sequence processed RNAs, the adapter ligation is performed using untreated total RNAs. In this case, the primary transcripts, whose 5' PPP group was not converted to 5' P, are not included in the library preparation (Figure 3D). After adapter ligation and eventually RNA fragmentation, the transcripts are retrotranscribed into cDNAs, which are subsequently amplified and sequenced.

The sequenced cDNAs, named reads, are mapped to the bacterial genome of reference. Comparison of the coverage profile (*i.e.* the number of reads mapped to a certain location of the genome) in the WT and Δ*rnase* strains was used to predict the location of the RNase processing positions[260]. For instance, when the processing of an endoRNase induces stabilization of one cleaved product and the degradation of the other one, the coverage profile upstream and downstream of the processing site differs between the WT and the Δ*rnase* strains (Figure 3D). In other studies, the 5' end coverage (*i.e.* the number of reads starting at a specific nucleotide) was also calculated. Comparison of the 5' end reads in the WT and Δ*rnase* strains allowed the identification of an endoRNase processing sites that are defined as the RNA 5' ends present in the WT strain, but not detected in the Δ*rnase* strain (Figure 3D). Unlike the total coverage analysis, this approach

pinpoints the processing positions also when the total coverage profile does not vary (*i.e.* both cleaved products are stable).

Whereas the mapping of the RNA 5' ends has been largely used to identify RNase cleavage positions, mapping of the RNA 3' ends was exploited only to investigate RNases that specifically produce RNA products with a 2', 3' -cyclic phosphate[262]. In other studies, mapping of the RNA 3' ends have been used to (i) detect premature terminated transcript due to riboswitches or attenuators[263], (ii) characterize the mechanisms of transcription termination[264] and (iii) investigate the differential RNA decay of mRNAs deriving from a polycistronic transcript[265].

Identification of endo and exoRNase specific cleavage positions in *S. pyogenes*

The above-described methods used to identify the endoRNase processing sites rely on the mapping of the RNA 5' ends. However, if after endoRNase processing, the RNA harboring the newly generated 5' end is degraded by exoRNases, the processing site will not be identified. To overcome this problem, the Identification of Specific Cleavage Position (ISCP) method was developed in our laboratory to pinpoint the endoRNase processing sites by comparing in the WT and Δ*rnase* strains the abundance of both RNA 5' and 3' ends[185] (Figure 4A). The RNAs were treated to allow adapter ligation at the RNA extremities and libraries containing either only primary or processed transcripts were generated and sequenced. The total 5' and 3' reads were calculated in both strains (schematically represented in Figure 4A). To precisely annotate an RNA end as an endoRNase cleavage position, the positions identified as more abundant in the WT strain than in the Δ*rnase* strain were filtered using different parameters including: (i) RNA expression threshold, (ii) cleavage ratio and (iii) proportion of the WT and Δ*rnase* cleavage ratios[185]. The mapping of the RNA 3' ends strengthens the identification of the processing sites as it allows to detect cleavage positions that would have not been uncovered with the solely analysis of the RNA 5' ends. ISCP was successfully applied to study RNase III in *S. pyogenes*. Mapping of only the RNA 5' ends was successively used in *E. coli* to also identify the RNase III processing sites[261].

The sequencing and mapping of the RNA 3' ends raise the possibility to also identify the RNA 3' ends produced by 3'-to-5' exoRNases, whose processing positions were never determined genome-wide at the nucleotide resolution. Therefore, the ISCP method was improved and used to pinpoint the direct targets of the 3'-to-5' exoRNases PNPase, RNase R and YhaM in *S. pyogenes*[204]. Libraries containing both primary and processed transcripts were generated from the WT strain and from each of the 3'-to-5' exoRNase deletion mutant strains (Δ*exornase*) (schematically represented in Figure 4B). The abundance of the RNA 5' and 3' ends was calculated at each nucleotide genome-wide and statistical analyses were performed to define the differentially

expressed ends, which were then filtered with the same parameters used in the RNase III study[185]. The RNA 3' ends corresponding to the positions where a 3'-to-5' exoRNase starts and stops the degradation of an RNA molecule were identified (Figure 4B), thereby giving insight into the 3'-to-5' exoRNase activity in *vivo*, in the human pathogen *S. pyogenes*[204].

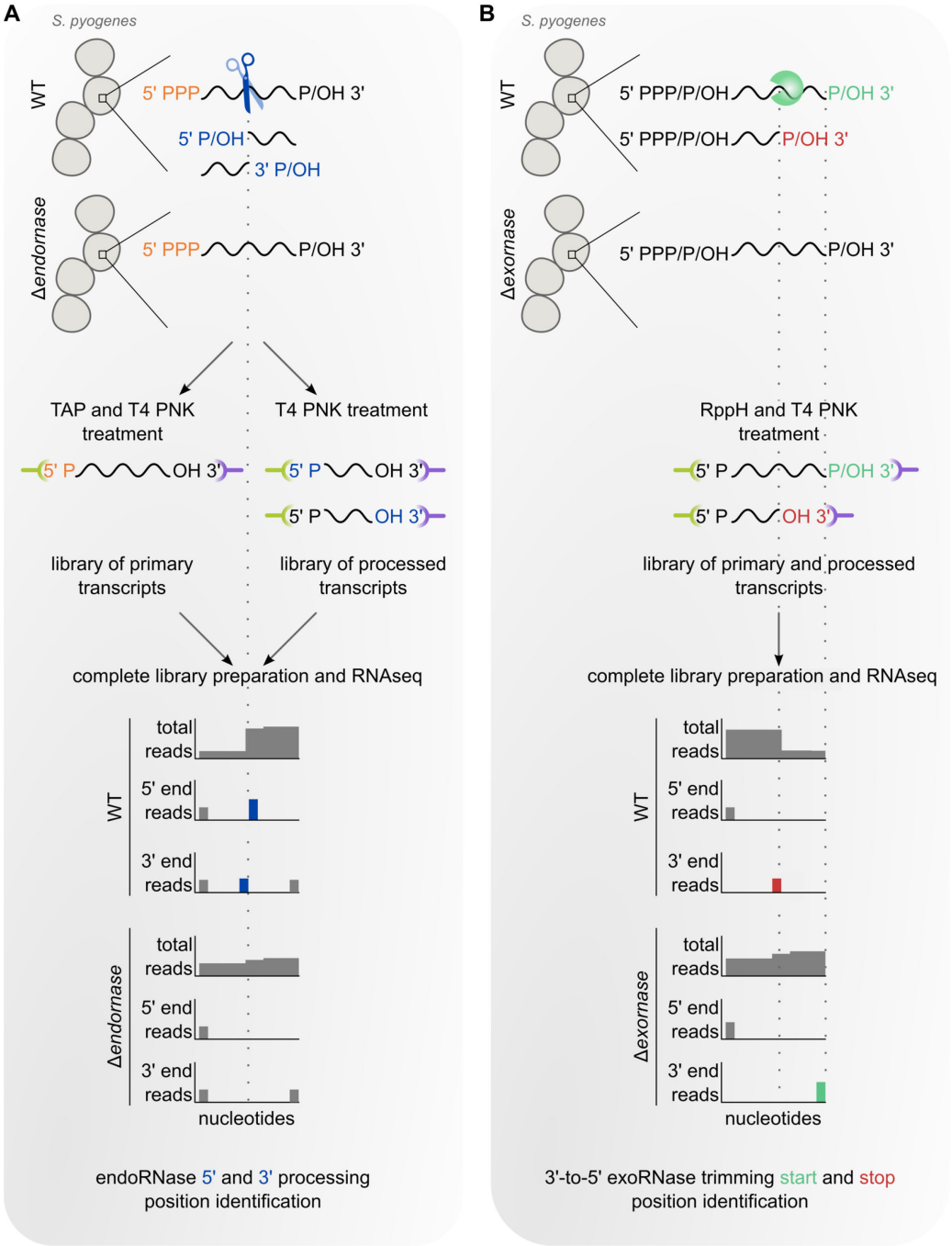


***Figure 4.* Analysis of RNA sequencing data to study RNases in *S. pyogenes*.**
A and **B.** Schematic representation of the strategy used in *S. pyogenes* to identify RNase processing sites by mapping the RNA ends in the WT strain and in the deletion mutant of an endoRNase (Δ*endornase*)[185] or exoRNase (Δ*exornase*) strains[204]. **A.** An endoRNase (blue scissors) cleaves a primary transcript (with a 5' triphosphate group, 5' PPP) and produces new RNA 5' and 3' ends (indicated in blue), both harboring either a monophosphate (P) or a hydroxyl group (OH) group. Libraries containing either only primary or only processed transcripts were generated by treating the RNAs in the presence or in the absence of Tobacco

acid pyrophosphatase (TAP). **B.** A 3'-to-5' exoRNase (green "pacman" symbol) degrades incompletely a transcript, producing a new RNA 3' end with either a P or OH group. Library containing both primary and processed RNAs were generated by treating the RNAs with pyrophosphohydrolase (RppH). **A** and **B.** T4 polynucleotide kinase (T4 PNK) is used to convert the OH in P. Adapters (in green and purple) are ligated to the RNAs with 5' P and 3' OH RNA ends. After the library preparation and RNA sequencing (RNAseq) the total, 5' and 3' coverages are calculated and here schematically represented. endoRNase cleavage site positions are identified as the RNA ends more abundant in the WT strain than in Δ*endornase* strain. 3'-to-5' exoRNase processing positions are identified as RNA 3' ends more abundant in the WT strain (stop positions) or more abundant in the Δ*exornase* strain (start positions).

1.3 The human pathogen *Streptococcus pyogenes*

S. pyogenes, also referred as group A streptococcus (GAS), is a chain forming, rod-shaped, Gram-positive bacterium, belonging to the Firmicutes phylum, Bacilli class and Lactobacillales order. One major phenotype observable while growing *S. pyogenes* on red blood agar plates is its ability to completely lyse erythrocytes and therefore to generate a halo surrounding the bacterial colonies. More than 200 serotypes of this bacterium have been identified via sequencing of the *emm* locus, which encodes the M protein, a major virulence factor located on the bacterial surface[266,267].

1.3.1 Pathogenesis

S. pyogenes is a strict human pathogen and while it asymptomatically colonizes the nasopharyngeal mucosa and the skin, it is also able to cause pathologies with a wide range of severity. Disorders caused by *S. pyogenes* consist mainly in superficial infections of the skin and throat epithelia leading to impetigo, pharyngitis and ultimately scarlet fever as a consequence of pharyngeal infection[268,269]. *S. pyogenes* can also penetrate deeper into the sterile tissues, for example infecting the epidermis and the subcutaneous tissue causing erysipelas and cellulitis, respectively[270]. The spreading of the bacteria in the soft tissue and infection of the *fascia* is responsible for necrotizing fasciitis, resulting in a severe tissue destruction. This is the reason why *S. pyogenes* is often referred to as a "flesh-eating bacteria". Further dispersion of the bacteria into the bloodstream can lead to bacteremia and subsequent streptococcal toxic shock syndrome, which are invasive diseases associated with high morbidity and mortality[271–274]. Complications due to *S. pyogenes*, after clearance of the infection, can be caused by reaction of the immune system to the bacterial antigens. These so-called post-infection sequelae include, for instance, the acute rheumatic fever and post-streptococcal glomerulonephritis[268,275].

1.3.2 Transmission

As the nasopharyngeal mucosa and the skin represent the major entry route of this bacterium, *S. pyogenes* is usually transmitted through respiratory drops and skin contact of infected

subjects. Transmission usually occurs from a human to another, however, infections caused by streptococcal contaminated food have been also reported[268]. Outbreaks due to *S. pyogenes* have been described as well and usually they are caused by strains that underwent genomic DNA modification with acquisition of genes encoding virulence factors[268].

1.3.3 Epidemiology and treatments

An epidemiological study in 2005 estimated that *S. pyogenes* was responsible for more than 600 million cases of pharyngitis and 18 million cases of severe infection yearly[271,276]. Despite the extensive usage of β-lactams antibiotic (*e.g.* penicillin) for the treatment of *S. pyogenes* infections, antibiotic resistance has not been observed; therefore, penicillin remains the treatment of choice[277]. In cases of allergies to penicillin or of treatment failure, due for example to *S. pyogenes* persistence inside the infected cells, macrolides (*e.g.* erythromycin) are used. However, *S. pyogenes* resistance to this class of antibiotics is increasing, thus a restricting usage is recommended[277,278]. Currently, some attempts to develop a vaccine to prevent *S. pyogenes* infection have been unsuccessful, likely due to the wide strain diversity[279,280].

1.3.4 The repertoire of virulence factors

An arsenal of virulence factors is produced by *S. pyogenes* allowing invasion and survival within the host, and defense mechanism against the host immune system are summarized in Figure 5. These factors exert specific functions during infection and therefore their expression is tightly regulated.

Adhesion

Successful *S. pyogenes* colonization relies on the adhesion to epithelial surfaces, for instance at the skin or oropharynx. The attachment to the host tissues is first mediated by weak interactions to collagen or the cell surface through pili and lipoteichoic acid (Figure 5A). Successively, stronger interactions occur through adhesins, including fibronectin binding proteins and also the aforementioned M protein, which can interact with cell surface glycosaminoglycans conferring tropism for specific tissues (Figure 5A)[268,281–283].

Dissemination

S. pyogenes produces a plethora of secreted factors to remodel the tissue and facilitate the spreading of the bacteria within the host. Streptokinase (Ska) and different deoxyribonucleases (DNases) are involved in the dissolution of fibrin clots and neutrophil extracellular traps,

respectively (Figure 5B). In addition, extracellular matrix component and hyaluronic acid are degraded by the Streptococcal pyrogenic exotoxin B (SpeB) and hyaluronidases, leading to disruption of the connective tissue[284,285] (Figure 5B).

Immune evasion

In order to survive within the host, *S. pyogenes* employs different strategies to antagonize the activity of the immune system: (i) inhibition of the complement system through the Streptococcal inhibitor of complement (SIC) and the ScpA peptides, (ii) degradation of immunoglobulins G (IgG) by SpeB and Ig degrading enzyme (IdeS) (Figure 5C). In addition, a hyaluronic acid capsule layer provides protection against recognition from both antibodies and complement system[268,283] (Figure 5C).

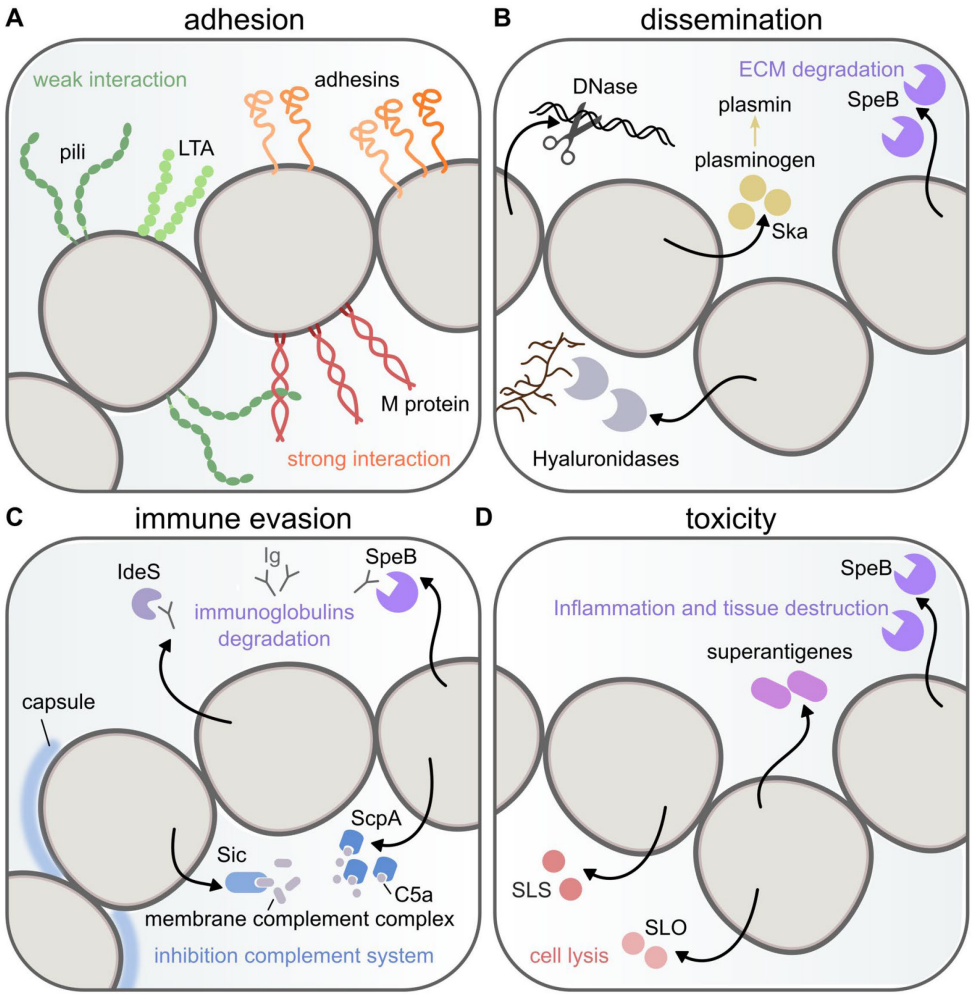


***Figure 5.* Virulence factors produced by *S. pyogenes*.**
Schematic representation of some of the main *S. pyogenes* virulence factors involved in adhesion, dissemination, immune evasion and toxicity. **A.** Pili and lipoteichoic acid (LTA) mediate weak interactions with the host cells, while the M protein and different types of adhesins are responsible for more stable interaction with the host tissue. **B.** Bacterial dissemination in the host is due to the activity of deoxyribonucleases (DNases), hyaluronidases, Streptokinase (Ska) and the Streptococcal pyrogenic exotoxin

B (SpeB) **C.** Immune evasion is accomplished through the activity of the Streptococcal inhibitor of complement (SIC) that blocks the membrane protein complex and through ScpA, a C5a peptidase, that degrades the complement-derived chemotaxin C5a. Disarming of the immune system occurs also through degradation of immunoglobulins (Ig) and by SpeB and Immunoglobulins degrading enzyme (IdeS). The capsule layer protects *S. pyogenes* from both Ig and complement system. **D.** Toxicity is due to the secretion SpeB, superantigens and lysis inducing proteins such as Streptolysin O (SLO) and Streptolysin S (SLS).

Toxicity

S. pyogenes produces several extracellular toxins (Figure 5D): (i) Streptolysin O (SLO) forming pore in the host cell membranes, (ii) Streptolysin S (SLS) inducing erythrocyte lysis, (iii) the SpeB protease, responsible for tissue damage and (iv) pyrogenic exotoxins (*e.g.* superantigens), which act as potent immunostimulatory molecules are responsible for severe diseases[275,286,287].

1.3.5 Streptococcal pyrogenic exotoxin B (SpeB) and its role in virulence

SpeB is a cysteine protease able to degrade a wide range of proteins and it is involved in multiple steps of *S. pyogenes* infection and pathogenesis (Figure 5). The gene encoding for SpeB (*speB*) is highly conserved and found in all the group A streptococci[288,289]. It was initially classified as a superantigen, but later demonstrated not to be involved in the massive stimulation of the immune system (typical of superantigens), but rather to function as a virulence factor due to its proteolytic activity[285,287]. To avoid targeting of bacterial cytoplasmic proteins, SpeB is produced as an inactive enzyme of 40 kilodalton (kDa) named SpeB zymogen (SpeBz) that undergoes multiple maturation events to generate a 28 kDa catalytically active protease defined as SpeB mature protein (SpeBm)[290]. The maturation occurs through intramolecular and, more frequently, intermolecular autocatalytic cleavage processes resulting in the elimination of a portion of the protein (named pro-domain), which prevents the SpeB protease activity[287,291,292]. In addition, reduction of a cysteine located in the active site is also required for SpeB activity[291,293].

Upon secretion and activation, SpeB degrades a large variety of host proteins from structural tissue components (*e.g.* fibrinogen) to molecules involved in the immune response (*e.g.* immunoglobulins, cytokines and chemokine)[294]. SpeB displays pro-inflammatory properties by cleaving and activating the precursor of interleukin-1β (IL-1β), a potent inducer of the antimicrobial immune response[295,296]. SpeB targets also streptococcal proteins that are located on the bacterial surface (*e.g.* M protein) and secreted virulence factors (*e.g.* Ska and C5a peptidase, DNase and IdeS), thereby remodeling the bacterial secretome[294]. In light of the diversity within the SpeB targets, it is clear that SpeB activity leads to apparently contrasting outcomes acting both as an anti-inflammatory molecule and promoting inflammation. However, it must be pointed out that in order to study the impact of SpeB on virulence, the bacterial location in the host, the levels of

speB expression and the activity of the protein during infection have to be considered. Moreover, many of the SpeB targets described have been investigated *in vitro* and few were confirmed *in vivo*, thereby considerably limiting the translational significance of these studies[294]. During the course of infection, *speB* expression decreases due to spontaneous mutations in the *covRS* locus encoding a two-component system (CovRS), which is one of the major inhibitors of *speB* transcription[297,298]. Reduced levels of SpeB *in vivo* correlates with progression of the infection, as many streptococcal virulence factors do not undergo proteolytic degradation[283,299]. Under these conditions, the SpeB-mediated activation of IL-1β is also reduced, leading to a decrease in the inflammatory response. It has been proposed that pathogenic bacteria evade the immune system by blocking the establishment of an inflammatory stage, which would impede bacterial invasion by diminishing the SpeB levels[296]. Another report showed, instead, that the levels of SpeB are not reduced in the majority of the strains derived from infected humans, independently of the type of diseases or bacterial serotype. These results indicate that a further clarification of the multiple roles of SpeB during an invasive disease is needed[300].

1.3.6 SpeB regulation

The expression of *speB* is regulated by multiple factors and environmental conditions (mainly investigated *in vitro*). Different expression levels according to the particular serotype studied were reported[290]. The *speB* gene is transcribed from two promoters (P and P1) and a third one (P2) was proposed, but likely was mis-annotated[161,301] (Figure 6). The expression of *speB* increases over time in cultures and it peaks during stationary phase of growth[288]. Environmental conditions such as variations in pH and salt concentration can also affect *speB* expression[302].

speB transcriptional regulation

A plethora of transcriptional regulator have been described to affect *speB* expression, but only few were proved to directly act on the *speB* promoter (Table 1).

speB regulators	Function	Activity	References
Direct transcriptional regulators			
RopB	Transcription factor	Activator	301,303–305
CovRS	Two-component system	Repressor	233,306–309
CcpA	Transcription factor	Activator	310
Indirect regulators at transcriptional level			
LacD.1	Aldolase	Repressor (via RopB)	311
N-terminal Vfr	Secreted peptide	Repressor (via RopB)	312,313
SIP	Secreted peptide	Activator (via RopB)	314,315
PepO	Endopeptidase	Repressor (proposed via RopB)	316
Transcriptional regulators with unknown mechanism			

speB regulators	Function	Activity	References
CpsY	Transcription factor	Repressor	317
CodY	Transcription factor	Activator	318
MgA	Transcription factor	Activator	319
LuxS	Transcription factor	Activator	320
Srv	Transcription factor	Repressor	321
SagP	Transcription factor	Activator	322
RofA	Transcription factor	Repressor	323

Table 1. **Transcriptional regulators of *speB* expression.**
List of *speB* direct transcriptional regulators (*i.e.* acting at the *speB* promoter), indirect regulators at transcriptional level (*i.e.* factors that control *speB* expression through the activity of direct transcriptional regulators) and transcriptional regulators whose mechanism of *speB* expression regulation is not deciphered.

One of the most studied *speB* transcriptional regulator is the Regulator of protease B (RopB), which binds inverted repeat regions upstream of the *speB* promoters and activates *speB* expression[301,303,304] (Figure 6). The gene encoding for this regulator (*ropB*) is located adjacent to *speB* and the two genes are divergently transcribed and separated by a long intergenic region, which encodes several small ORFs, including the SpeB inducing peptide (SIP)[314,315] (Figure 6). SIP consists of a small peptide whose expression is induced at high cell density conditions during which SIP is secreted and subsequently reimported into the bacteria where it binds RopB. This interaction enhances the ability of RopB to bind DNA and therefore it promotes *speB* expression activation[315] (Figure 6). It is believed that the RopB-mediated regulation of *speB* transcription is responsible for the increase of *speB* expression during growth. The activity of RopB is inhibited by the tagatose 1-6 bisphospate aldolase (LacD.1)[311] and the Virulence factor related (Vfr)[313] (Figure 6). LacD.1 binds RopB, possibly in accordance to the nutritional status of the cell and consequently prevents *speB* expression activation[305]. Vfr harbors an amino-terminal (N-terminal) secretion signal system, which is cleaved, secreted and internalized[312] (Figure 6). The Vfr N-terminus directly interacts with RopB thereby preventing RopB binding upstream of the *speB* promoter[313].

As mentioned above, the CovRS two-component system directly represses *speB* expression, in particular the response regulator CovR binds the *speB* promoter and inhibits transcription[307,308]. This system senses external signals and stimuli through the sensor kinase CovS, which in turn phosphorylates or dephosphorylates the transcriptional regular CovR. While phosphorylation enhances the CovR-mediated repression of some target genes, dephosphorylation has an opposite effect on another subset of genes[309]. It has been observed that CovS attenuates the CovR-mediated downregulation of *speB* expression, indicating that phosphorylated-CovR does not repress *speB* expression[309,324] (Figure 6).

The last confirmed direct regulator of *speB* is the Catabolite control protein (CcpA), which activates *speB* expression during nutrient-limiting conditions[325,326]. Although direct interaction of

CcpA with the DNA region in proximity of *speB* promoter was demonstrated, the exact binding site has not been clarified yet[325,326].

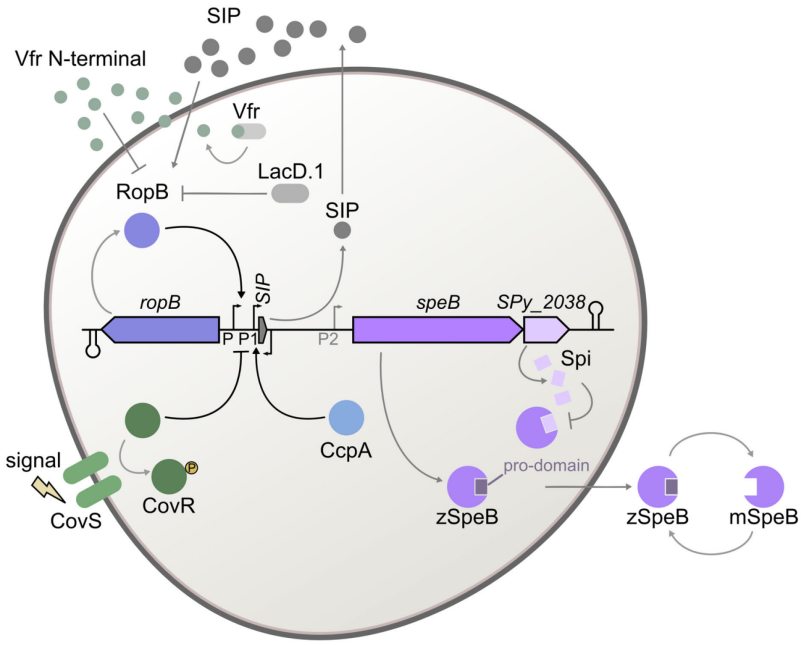


Figure 6. **Regulation of SpeB expression and activity.**
Schematic representation of some well-studied regulatory mechanisms of SpeB expression and activity. The *speB* locus is shown with the predicted promoters (P, P1 and P2) and terminators. The *speB* gene is transcribed in the same direction than *SPy_2038* (encoding Spi, an inhibitor of SpeB protease activity) and in the opposite direction than *ropB* (encoding RopB). Three direct *speB* transcriptional regulators have been described: RopB, CovR and CcpA. Their positive or negative effects on *speB* expression are indicated. RopB-mediated regulation of *speB* expression is affected by three factors: SIP, LacD.1 and Vfr. SIP. SpeB is produced as inactive zymogen (zSpeB), which is secreted and then activated through an autocatalytic process resulting in the removal of the pro-domain and generation of the mature SpeB (mSpeB). mSpeB processes the zSpeB producing more active SpeB proteases. Spi, mimicking the pro-domain, can block the protease activity intracellularly.

speB post-transcriptional regulation

The post-transcriptional regulation of *speB* expression has remained unexplored. *speB* transcript abundance and production of the active SpeB were shown to be affected by the pleiotropic effect locus (*pel*), which codes for both the hemolysin SLS and a regulatory RNA element[327,328]. The mechanism of *speB* regulation by *pel* RNA is undeciphered and it is unclear whether it is due to direct or indirect effects. *speB* transcript stability was shown to increase as a consequence of the medium acidification during growth[329]. Two *speB* transcript isoforms were identified in M49 *S. pyogenes* strain NZ131 and RNase Y was shown to control the stability of the shorter *speB* isoform[329]. In addition, the abundance of *speB* strongly decreases in the absence of RNase Y, but the mechanism of RNase Y-mediated regulation of *speB* has not been characterized[168,329] and it will be further investigated in this thesis (see section 3.1).

SpeB post-translational regulation

At the post-translational level, modulation of SpeB protease activity occurs at first through the production of the active mature SpeB (mSpeB) by removal of the pro-domain (Figure 6). Two peptidyl-prolyl cis-trans isomerases (RopA and PrsA) ensures the correct maturation of the inactive zSpeB[330,331]. A second control mechanism relies on the SpeB inhibitor (Spi), which mimics the SpeB pro-domain and therefore blocks the SpeB protease activity (Figure 6). Spi was shown not to be secreted. Therefore it must inhibit SpeB intracellularly to prevent cytosolic protein degradation in the case of premature SpeB activation[332]. Both genes encoding Spi and PrsA (SPy_2038 and *prsA*, respectively) are co-transcribed with *speB*[331].

1.3.7 Regulation of gene expression in *S. pyogenes*

S. pyogenes tightly controls gene expression through a variety of regulatory systems; transcriptional regulators have been largely studied while little is known about post-transcriptional regulatory mechanisms.

Transcriptional regulation

The integration of external stimuli is often achieved through the activity of stand-alone response regulators and two-component systems. The best characterized stand-alone response regulators are the aforementioned RopB, Multiple gene activator (Mga) and RofA like protein (RALP), which can respond to quorum sensing signals and nutrient availability[317]. These three regulators are specialized in the control of a subset of virulence genes and at specific growth phases. For instance, MgA positively regulates genes involved in the adhesion and immune evasion during exponential growth phase and RopB controls the expression of secreted proteins, including proteases, during late growth phase[333,334]. Thirteen two-component systems have been identified, but most of them have not been characterized in detail[317], with the notable exception of the CovRS system, a master regulator that affects the expression of 15% of *S. pyogenes* genome[335].

Under specific environmental conditions, bacteria often use alternative σ factors to direct the RNA polymerase at specific promoter regions. *S. pyogenes* does not extensively use alternative σ factors to modulate gene expression in response to environmental conditions. Instead, this is achieved by the network of stand-alone response regulators and two-component systems[317]. However, an alternative σ factor (σ^X) was identified in *S. pyogenes* and it was shown that this σ factor is poorly expressed under standard growth conditions in nutrient-rich medium[336]. It is known that σ^X can induce the expression of genes involved in competence acquisition[337,338].

Post-transcriptional regulation

The role of few sRNAs has been characterized in *S. pyogenes* [68] and one of the most studied is the *trans*-acting sRNA FasX. Besides positively regulating *ska* mRNA, as already described[233] (Figure 2D), FasX also controls the translation of the pili operon and of *prtF1-2* mRNA (encoding for two fibronectin-binding proteins). FasX interacts with the 5' end region of these mRNAs and inhibits their translation[339–341]. Hence, based on the observation that FasX promotes the streptokinase production and inhibits the synthesis of proteins involved in the bacterial adherence with the host (*i.e.* pili and fibronectin-binding proteins), this sRNA is considered to play a role in the switch from bacterial adherence to dissemination during infection.

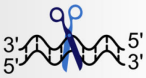
Activity[a]	RNases[b]	Gene[c]	Phenotype[d]	Functions[e]
ssRNA endoRNases 	RNase Y	*SPy_1633* (*rny*)	growth delay; decreased virulence	RNA decay/ maturation
	YbeY	*SPy_0473* (*ybey*)	-	rRNA maturation
	YacP/RaeI	*SPy_0747* (*spnA*)*	-	RNA decay
	YhcR	*SPy_1937*	-	utilization nucleic acids*
	RNase Z	*SPy_0924* (*rnz*)	-	tRNA maturation
	RNase P	*SPy_0246* (*rnpA*)	-	tRNA maturation
	RNase HII	*SPy_1162* (*rnhB*)	-	RNA replication
	RNase HIII	*SPy_1841* (*rnhC*)	-	RNA replication
dsRNA endoRNases 	RNase III	*SPy_0531* (*rnc*)	no growth defects	rRNA and CRISPR maturation
	Mini III	*SPy_1940* (*mrnc*)	no growth defects	rRNA maturation
	RNase M5	*SPy_0261* (*rmnV*)	-	rRNA maturation
5'-to-3' exoRNases  endoRNases?	RNase J1	*SPy_1876* (*rnjA*)	lethal	RNA decay; rRNA maturation
	RNase J2	*SPy_1020* (*rnjB*)	lethal	RNA decay
3'-to-5' exoRNases 	PNPase	*SPy_1946* (*pnpA*)	growth delay	RNA decay
	YhaM	*SPy_0267* (*yhaM*)	defect in cold adaptation	RNA decay
	RNase R	*SPy_0503* (*rnr*)	defect in cold adaptation	rRNA maturation; RNA decay
	NrnA*	*SPy_0720*	-	oligoRNA degradation

Table 2. **RNases present in *S. pyogenes*.**
The RNases identified in *S. pyogenes* are listed and separated based on their activity. [a] Schematic representation of an RNA molecule cleaved by an endoribonuclease (endoRNase), depicted with scissors and of an RNA molecule degraded by 5'-to-3' or 3'-to-5' exoribonuclease (5'-to-3' or 3'-to-5' exoRNase) depicted with orange and green "pacman" symbols, respectively. ssRNA: single-stranded RNA; dsRNA: double-stranded RNA. RNase J1 likely acts more as 5'-to-3' exoRNase then endoRNase *in vivo*, RNase J2 activity is unclear. [b] NrnA (NanoRNase A) was shown to act also as a 5'-to-3' exoRNase. Of note, *S. pyogenes* does not encode for NanoRNase B (NrnB). [c] Gene names are based on the NCBI and UniProt databases. Previous searches failed to identify YacP/RaeI endoRNase in *S. pyogenes*[84]. SpnA could be a homolog of YacP/RaeI (identity in BLAST: 43.84%). [d] The phenotypes observable when the gene encoding an RNase is deleted are reported for RNase Y[168,329], RNase III[185], Mini III (unpublished data), RNases J1 and J2[195], PNPase[204], YhaM[204] and RNase R[204]. The minus sign indicates that the effect of the RNase deletion is not known in *S. pyogenes*. [e] The

RNase functions in grey are deduced from the *B. subtilis* homologs. YhcR was shown to be an extracellular not specific endonuclease targeting DNA and RNA.

Activity[a]	Helicases[b]	Gene[c]	Homolog *B. subtilis*[d]	Predicted functions[e]
DEAD-box RNA helicases 	CshB	*SPy_1659*	CshB	putative ATP-dependent RNA helicase
	DeaD	*SPy_1415*	CshA (YdbR)	putative ATP-dependent RNA helicase
	DeaD2	*SPy_1369*	YfmL	putative ATP-dependent RNA helicase
	-	*SPy_0669*	-	putative DEAD/DEAH box helicase, phage associated

Table 3. **Helicases annotated in *S. pyogenes*.**
List of putative helicases belonging to the DEAD-box family present in *S. pyogenes*. [a] Schematic representation of a DEAD-box RNA helicase (composed of two domains) associated to ATP and binding a structured RNA. [b,c] The name of the gene encoding the helicase and the names of each helicase are based on the NCBI and UniProt databases; the minus sign indicates there is not name associated. [d] Homologs in *B. subtilis* are identified by BLAST search tool. [e] The predicted function of the *S. pyogenes* gene from NCBI is reported.

A limited number of studies investigated the impact of RNases in post-transcriptional regulation in this bacterium and the role of helicases in RNA metabolism has not been explored yet. RNases and putative annotated helicases present in *S. pyogenes* are listed in Table 2 and Table 3, respectively. The features and functions of the RNases from *S. pyogenes* have been already described in sections 1.1.3 and 1.1.4). Knowledge on RNases in *S. pyogenes*, prior the studies from our laboratory aiming at identifying RNase III, PNPase, RNase R and YhaM processing positions[185,204] (see sections 1.1.4 and 1.2.2) was restricted to a limited number of putative direct targets[53,195,329,342]. RNA stability was shown to play an important role in the regulation of gene expression in *S. pyogenes*. In this bacterium, mRNAs are grouped in two classes (I and II), based on their detection either during exponential or stationary phases of growth, respectively[195,342,343]. Class I transcripts are present only during exponential growth phase, since they are the preferred substrate of RNases J1 and J2, and therefore they are processed by these RNases and rapidly degraded. On the other side, class II transcripts, including the ones encoding virulence factors (*e.g.* SLS and the DNase Sda1), are still detectable during stationary growth phase. This is due to the fact that the degradation initiation of these transcripts by RNases J1 and J2 is delayed compared to class I mRNAs. PNPase was shown to be involved in the degradation of transcripts from both classes, after RNases J1 and J2 processing[342,343].

Transcriptomic studies were performed in *S. pyogenes* to investigate the impact of RNase Y on transcript abundance[168] and stability[106]. The knowledge of RNase Y activity and direct targets is limited to few examples in *S. pyogenes*. It is known that RNase Y likely processes *ropB* mRNA that it controls the expression of virulence factors, including *speB*, but the mechanism is not

characterized[168,329]. Overall, knowledge on RNase Y direct targets and requirements for processing has remained limited in this bacterium compared to other Gram-positive bacteria, such as *B. subtilis* and *S. aureus*. Identification and characterization of the RNase Y targets in *S. pyogenes* will be therefore one of the objects of the present thesis.

1.4 Aim of the thesis

RNase Y exhibits species-specific effects on gene expression, bacterial physiology and different strategies of target recognition, indicating that our understanding of RNase Y function is not complete yet. The specific properties of RNase Y orthologs highlight the importance of investigating the characteristics of this endoRNase not only in well characterized model microorganisms, but also in other bacteria. In particular, in light of the key role of RNase Y in the regulation of virulence genes, the role of RNase Y in bacterial pathogens should be further examined.

The aim of this thesis was to elucidate the characteristics and functions of RNase Y from the strict human pathogen *S. pyogenes* using different strategies. First, we identified the genome-wide RNase Y direct targets *in vivo* (*i.e.* targetome), using RNA sequencing. The mapping of the exact RNase Y processing positions gave insights into RNase Y target determinants. Second, we investigated the interplay between RNase Y and the 3'-to-5' exoRNases during RNA degradation, by comparing the RNase Y targetome with the 3'-to-5' exoRNase targetomes already available. Third, we further examined the role of RNase Y in the regulation of the major virulence factor Streptococcal pyrogenic exotoxin B (SpeB).

This study provided new and profound insights into the features of *S. pyogenes* RNase Y and expanded the current understanding of RNA degradation and dynamics in Gram-positive bacteria. The investigation of RNase Y-mediated regulation of *speB* expression may pave the way for further understanding of RNase Y role in virulence.

Materials & Methods

The methods used in the present thesis are based on published techniques that can be also found in previous publications from our laboratory[185,204,344]. The methods presented here were performed, if not specified, by myself (L.B.) or by myself together with Dr. Anaïs Le Rhun (A.L.R.) and Dr. Anne-Laure Lécrivain (A.-L.L.) and with the help of Dr. Thibaud T. Renault (T.T.R.) and Dr. Rina Ahmed-Begrich (R.A.-B.) for the bioinformatics and Karine Hahnke (K.H.) and Solange Materne (S.M.; master student in the Charpentier group supervised by L.B.) for technical work.

2.1 Bacterial culture

Table 4 lists all the bacterial strains used in this thesis. *S. pyogenes* SF370 (M1 GAS; WT strain) and the deletion mutant strains were cultured in THY containing: Todd Hewitt Broth (THB; Becton Dickinson) containing heart infusion (3.1 g/L), peptone (20 g/L), dextrose (2g/L), sodium chloride (NaCl, 2 g/L), sodium phosphate (Na_2HPO_4; 0.4 g/L) and sodium carbonate (Na_2CO_3; 2.5 g/L) and supplemented with 0.2% yeast extract (Servabacter®). The above-mentioned components were dissolved in distillated water (dH_20), the pH was adjusted at 7.0 and the medium was autoclaved. The growth of *S. pyogenes* strains was conducted in THY without shaking at 37°C and with 5% CO_2. The growth on solid medium was conducted on TSA blood agar (Trypticase™ Soy Agar, Becton Dickinson), which was supplemented with 3% defibrinated sheep blood (Oxoid). *E. coli* was grown at 37°C, shaking (180 rpm/min) and in Luria-Bertani (LB Broth Miller, Becton Dickinson) broth medium, consisting of tryptone (10 g/L), yeast extract (10 g/L) and sodium chloride (0.5g/L) dissolved in dH_20. For the growth of *E. coli* on solid medium, LB broth medium with the addition of agar (15g/L; Miller) was used. For the growth of *S. pyogenes,* the strains were grown over-day (at 37°C and with 5% CO_2) on a TSA blood plate streaked with the bacteria glycerol stocks stored at − 80°C. Subsequently the bacteria were inoculated in 10 ml (for RNA and protein extraction and growth curve analysis) or 2 ml (for the strains containing plasmids) of

medium for the overnight growth. The cultures were then diluted (depending on the experiment, see below) and the growth was monitored by measuring the optical density at 620 nanometers ($OD_{620\ nm}$) with a microplate reader (Biotek Epoch 2). At the end of each experiment the over-day culture was used to streak a TSA blood plate in order to assess the presence of contaminating bacteria. When needed, antibiotics were added to the media at the following final concentrations: 300 µg/ml kanamycin (Corning™) and 3 µg/ml erythromycin (Sigma-Aldrich) for *S. pyogenes*; 50 µg/ml kanamycin (Corning™) and 300 µg/ml erythromycin (Sigma-Aldrich) for *E. coli*. Prior addition to the media the kanamycin and the erythromycin were dissolved in Milli-Q H_20 and ethanol (EtOH), respectively and stored at − 20°C. When necessary, anhydrotetracycline hydrochloride (AHT) (Sigma-Aldrich) was supplemented to the medium, as previously described[195]. AHT was used at a final concentration of 0.1 ng/µL and was used to activate the conditional tetracycline-inducible promoter (P_{tet}).

Strain	Relevant characteristics	Source
Streptococcus pyogenes		
WT		
EC2224	SF370 (M1 serotype)	ATCC 700294
Δ*rny*		
EC2246	Δ*rny*::lox72	185
Δ*rnc*		
EC2249	Δ*rnc*::lox72	185
Δ*rnr*		
EC2254	Δ*rnr*::lox72	204
Δ*pnpA*		
EC2297	Δ*pnpA*::lox72	204
Δ*rny::rny*		
EC2298	Δ*rny*::lox72Δlox72::*rny*-TT3-lox72	344
Δ*yhaM*		
EC2347	Δ*yhaM*::lox71- P*ermAM/B-ermAM/B*-lox66	204
Δ*rnhB*		
EC2251	Δ*rnhB*::lox72	344
Δ*mrnc*		
EC2271	Δ*mrnc*::lox72	344
P_{tet}*rnjA*		
EC2353	P_{tet}*rnjA*	Laboratory collection
Δ*pnpA*Δ*rny*		
EC2389	Δ*pnpA*::lox72Δ*rny*::lox71- P*ermAM/B-ermAM/B*-lox66	204
Δ*rny*Δ*rnr*		
EC2310	Δ*rny*::lox72Δ*rnr*::lox72	204
Δ*yhaM*Δ*rny*		
EC2392	Δ*yhaM*::lox71-P*ermAM/B-ermAM/B*-lox66Δ*rny*::lox71-P*ermAM/B-ermAM/B*-lox66	204
Δ*speB*		
EC2356	Δ*speB*::lox72	344
Escherichia coli		
RDN204 Top10 (Host for cloning)	F– *mcrA* Δ(*mrr-hsd*RMS-*mcrBC*) φ80*lacZ*ΔM15 Δ*lacX*74 *recA1 araD*139 Δ(*ara-leu*)7697 *galU galK* λ– rpsL(StrR) *end*A1 *nupG*	Invitrogen

Table 4. **Strains used in the study.**
The name and genotype of the strains used in this thesis are reported. The deletion mutant strains were generated using the Cre-lox system[345]. The TT3 represents the transcriptional intrinsic terminator of the

bacteriophage T3[346], *ermAM/B* denotes the erythromycin resistance cassette and P*ermAM/B* the corresponding promoter. P_{tet} designates the tetracycline-inducible promoter, which controls the expression of *rnjA* (coding for RNase J1)[195,347,348].

2.1.1 *S. pyogenes* growth

The growth comparison of the *S. pyogenes* WT and deletion mutant strains was performed as previously described by our laboratory[185]. After diluting the *S. pyogenes* overnight cultures to an $OD_{620\ nm}$ of 0.2 in 5 ml of medium, the bacteria were centrifuged at 3200 g for 5 min at 20°C. The supernatant was discarded and the cell pellet was resuspended in 500 µl of medium, which was used to inoculate a 100 ml-flask containing 50 ml of THY supplemented with 0.2% yeast extract. The growth curves were performed in biological triplicates and the standard deviations are reported.

2.1.2 Bacterial transformation

Chemically competent *E. coli* cells were prepared according to a standard protocol[349] and the cells were then transformed using heat-shock protocol with some modifications[349]. Chemically competent *E. coli* were mixed with plasmid DNA (100 ng or 5 µl from the ligation reaction) and incubated for 10 min on ice. To cause the heat-shock, the cells were then incubated at 42°C for 1 min and subsequently on ice for 3 min. After addition of 750 µl of LB medium, the cells were recovered at 37°C for 1h or 1h and 30 min for plasmids with kanamycin or erythromycin resistance cassettes, respectively. *S. pyogenes* competent cells for deletion mutant generation were prepared as previously described[185,350]. *S. pyogenes* WT cells were grown, over-day, until $OD_{620\ nm}$ 0.25 in THY supplemented with 250 mM Sucrose (AppliChem) and 40 mM L-threonine (Sigma-Aldrich). Following centrifugation (3500 g for 10 min at 4°C), the cell pellets were washed twice with 5 ml of ice-cold 0.5 M sucrose and resuspended in 250 µL of 0.5 M sucrose and 20% glycerol. Before snap freezing with liquid nitrogen, the $OD_{620\ nm}$ of the generated competence cells was measured. Prior transformation, the *S. pyogenes* competent cells were defrosted on ice and the $OD_{620\ nm}$ was adjusted to 1. Subsequently the competent cells were incubated with approximately 10 µg of linearized plasmid for 1h on ice and electroporated in a 0.1 cm cuvette (Bio-Rad) with 1.5 kV, 400 Ω, 25 µF pulse. After a recovery growth of 2h in THY medium, the bacteria were spread on TSA-blood plates containing erythromycin.

For transformation of *S. pyogenes* (WT or deletion mutant strains) with plasmids, the cells were grown over-day in THY until $OD_{620\ nm}$ 0.3, incubated on ice for 10 min and washed twice with 25 ml and 12 ml of ice-cold Milli-Q H_20, respectively. Prior each washing step with ice-cold Milli-Q H_20, the cells were centrifuged at 3500 g for 10 min at 4°C. The cells were finally resuspended

in 6.25 ml of 20% glycerol and centrifuged at 3500 g for 10 min at 4°C. After resuspension in 1 ml of ice-cold sterile 20% glycerol, the $OD_{620\ nm}$ was measured and the cells were snap feezed with liquid nitrogen. After defrosting the MQ-competent cells on ice, 500 or 800 ng plasmids were incubated together with $OD_{620\ nm}$ 1 or 2 of competent cells for 1h on ice, respectively. The electroporation was conducted at 1.8 kV, 400 Ω, 25 µF. Following 2h of regeneration time in THY medium, the bacteria were grown on TSA-blood plates supplemented with kanamycin. All the plasmids used in the bacterial transformation in *E. coli* and *S. pyogenes* are listed in Table 5.

Plasmid	Relevant characteristics [a]	Source
Plasmids used for gene deletion in *S. pyogenes*		
pEC707(pUC19)	Fra^-, Meb^-, Km^R, Tc^R, Mc^R	New England Biolabs; 351
pEC85	*repDEG*-pAMß1, *aphIII*-Pjh1, ColE1	53
pEC454	pUC19Ωlox71-*ermAM/B*-lox66	344
pEC455	pEC85Ω*PgyrA-cre*	352
pEC801	pRO1600/ColE1, *ampR*	pSEVA141 De Lorenzo's lab
pEC2145	pEC801Ω*speB*::lox71-P*ermAM/B-ermAM/B*-lox66	This study
speB* ectopic expression in *S. pyogenes		
pEC2146	pEC85ΩP_{gyrA}*speB*	This study
pEC2249	pEC85ΩP_{gyrA}-*speB*(G−137A)	This study
pEC2250	pEC85ΩP_{gyrA}-*speB*(G−131A)	This study
pEC2263	pEC85ΩP_{gyrA}-*speB*(G−137A_G−131A)	This study
pEC2264	pEC85ΩP_{gyrA}-*speB*(Δ−147−121)	This study
pEC2265	pEC85ΩP_{gyrA}-*speB*(Δ−157−111)	This study
Luminescence assay in *S. pyogenes*		
pEC2173	pLZ12Km2-P23R:TA:*ffluc*	Gift from Thomas Proft (Addgene plasmid # 88900)
pEC2248	pEC2173Ω*PspeB*	This study

Table 5. **Plasmids used in the study.**
The name and the description of the plasmids used in this thesis are reported. [a] *repDEG*-pAMß1: origin of replication specific for Gram-positive bacteria[353]; *aphIII*: kanamycin resistance cassette for selection in both *E. coli* and *S. pyogenes* [354]; ColE1: replication sequence specific for *E. coli* [354]; lox71 and lox66: sequences recognized by the Cre recombinase[345]; pRO1600/ColE1: combined replication sequence of pRO1600 and ColE1[355]; *ermAM/B*: erythromycin resistance cassette; *ampR*: ampicillin resistance cassette; P_{gyrA}: constitutive promoter of *gyrA* from *Streptococcus agalactiae*; P_{speB}: promoter region of the *speB* gene, the region included is described in section 2.2.2; pLZ12Km2: vector backbone containing kanamycin resistance cassette; P23: lactococcal constitutive promoter P23[356]; *ffluc*: gene encoding the firefly luciferase[357]. For the plasmids used in the *speB* ectopic expression experiments the positions of the guanosine (G) substituted into adenosine (A) or of the deleted regions (Δ) are indicated relative to the *speB* start codon.

2.2 DNA manipulation and cloning

All the plasmids (Table 5) were purified from *E. coli* TOP 10 cells after an overnight growth in 5, 10 or 20 ml of LB medium, which contained the respective antibiotic for selection, using QIAprep Spin MiniPrep Kit (Qiagen) following the manufacturer's instructions. Alternatively, the plasmids constructed for deletion mutant strain generation were extracted using the QIAprep Spin MidiPrep Kit (Qiagen), from 100 ml overnight cultures. *S. pyogenes* genomic DNA was extracted using the NucleoSpin® Microbial DNA kit (Macherey-Nagel) following the protocol for Gram-

positive and Gram-negative bacteria except for the lysis step, which was performed by agitation for 15 min using a vortex. For the purpose of cloning the oligonucleotides used are listed in Table 6. Restrictions enzymes (specified below at each specific section) were used to linearize the plasmids and the digestion was conducted at 37°C for at least 1h following the manufacturer's instructions. To avoid the relegation of the linearized plasmids, 1U of the FastAP Thermosensitive Alkaline Phosphatase (Thermo Scientific) was added to the reaction and incubated at 37°C for at least 10 min. If not specified otherwise, the polymerase chain reaction (PCR) products were obtained using the Phusion High Fidelity Polymerase (Thermo Scientific) and the primers are listed in Table 6. The PCR products were purified either directly from the PCR reaction with the QIAquick PCR Purification Kit (Qiagen) or by gel extraction with the QIAquick Gel Extraction Kit (Qiagen) after gel electrophoresis and following the manufacturer's instructions. The linearized plasmids and the PCR inserts were ligated using 1U of T4 DNA ligase (Thermo Scientific), following the manufacturer's instructions. The sequencing of the PCR-generated DNA fragments and of the constructed plasmids was performed at Microsynth, Switzerland, using oligonucleotides listed in Table 6.

Oligo	Sequence 5′-3′ [a]	F/R [b]	Usage [c]	Target [d]
Δ*speB*				
OLEC7565	AAAGGATCC*ATGTCAAAAATACGTTACGCATG*	F	Cloning	Up fragment (pEC2145)
OLEC7566	TATAATGTATGCTATACGAACGGTA*TTTTTTTATACCTCTTT CAAAATAAGTTAATCTAC*	R	LM-PCR	
OLEC7902	ATAGCATACATTATACGAACGGTA*GACGGACGTAACTTCT ACCATGTT*	F	LM-PCR	Dw fragment (pEC2145)
OLEC7569	*AAAGGATCCTGTTGTGTGATGATTGACAAGCTG*	R	Cloning	
OLEC7563	*TGAATGCCTAATGAATTCAACGG*	F	PCR, SEQ	Upstream *speB*
OLEC7570	*GTGTTTTTGGTCTCATTGTAGAAGT*	R	PCR, SEQ	Downstream *speB*
pEC2146				
OLEC7968	CCTTTCTAGA*CTATCATTTTCAATGAAAGAAGTCACTAATA AAATGTGA*	F	Cloning	P*gyrA* (pEC455)
OLEC7969	CATAGTAGGCGCCTCCTTTTAACCTTAT*TACATTGTACCAT AATTTAGGTAAAATTGCGATGAT*	R	LM-PCR	
OLEC7970	ATCATCGCAATTTTACCTAAATTATGGTACAATGTA*ATAAG GTTAAAAGGAGGCGCCTACTATG*	F	LM-PCR	*speB* CDS
OLEC7971	CCCAGAATTC*CTAAGGTTTGATGCCTACAACAGCAC*	R	Cloning	
pEC2249				
OLEC8388	*GTCAACTAACCGTATTATTGTCTATTACCAT*	F	TS-PCR	*speB* 5′ UTR (pEC2146)
OLEC8389	*GTCAACTAACCGTATTATTGTCTATTACCAT*	R	TS-PCR	
pEC2250				
OLEC8390	*GTCAACTAACCGTGTTATTATCTATTACCAT*	F	TS-PCR	*speB* 5′ UTR (pEC2146)
OLEC8391	*ATGGTAATAGATAATAACACGGTTAGTTGAC*	R	TS-PCR	
pEC2263				
OLEC8392	*GTCAACTAACCGTATTATTATCTATTACCAT*	F	TS-PCR	*speB* 5′ UTR (pEC2146)
OLEC8393	*ATGGTAATAGATAATAATACGGTTAGTTGAC*	R	TS-PCR	
pEC2264				
OLEC8394	*GTTGGGTTGTCAGTGTCATCATGGTATCAGCGACAT*	F	TS-PCR	*speB* 5′ UTR (pEC2146)
OLEC8395	*ATGTCGCTGATACCATGATGACACTGACAACCCAAC*	R	TS-PCR	

Oligo	Sequence 5′-3′ [a]	F/R [b]	Usage [c]	Target [d]
pEC2265				
OLEC8396	*GAATAATTGGGTTGGGTTAGCGACATCGTATGATAA*	F	TS-PCR	*speB* 5′ UTR (pEC2146)
OLEC8397	*TTATCATACGATGTCGCTAACCCAACCCAATTATTC*	R	TS-PCR	
pEC2248				
OLEC8386	CGAGCTC*ATGTCAAGCCTTCCTAGTTGATGTCA*	F	Cloning	*speB* 5′ UTR
OLEC8387	TACCCGCGG*TGGCTATATCATAGCTGCTTATTTTGCT*	R	Cloning	
Sequencing				
OliRN228	*GGAACGAAAACTCACGTTAA*	F	SEQ	pEC85
OLEC787	*TGTGGTTACGTGGTTTTTAAC*	R	SEQ	
OLEC3224	*TGTAAAACGACGGCCAGT*	F	SEQ	pEC707 pEC2173
OLEC3225	*CAGGAAACAGCTATGACC*	R	SEQ	
OLEC3600	*CCAGGGTTTTCCCAGTCACGAC*	F	SEQ	pEC801
OLEC3590	*AGCGGATAACAATTTCACACAGGA*	R	SEQ	
OLEC1938	*TCAATCGAGAATATCGTCAACTGTTTACTAAA*	F	SEQ	*ermAM/B*
OLEC1937	*TTGCTGTTTCGATTTTTATGATATGGTGC*	R	SEQ	
OLEC5336	*GGGGGATGTGCTGCAAGGCG*	F	SEQ	pEC802
OLEC5337	*TCCGGCTCCTATGTTGTGTGG*	R	SEQ	

Table 6. **Oligonucleotide primers used for cloning and sequencing.**
The name, the sequence, the directionality, the function and the respective targets of the oligonucleotides used in this thesis are reported. [a] *italic*: sequence annealing to the template; underlined: restriction site; [b] F: forward primer; R: reverse primer; [c] LM-PCR: ligation-mediated PCR; TS-PCR: two-stage PCR; SEQ: sequencing; [d] 5' UTR: 5' untranslated region.

2.2.1 Construction of gene deletion strains

Chromosomal deletion of *speB* from *S. pyogenes* genome was performed using the Cre-Lox system. Regions upstream and downstream of the *speB* gene were amplified by PCR from WT genomic DNA with Phusion polymerase (Thermo Fisher) using OLEC7565/OLEC7566 (upstream fragment) and OLEC7569/OLEC7902 (downstream fragment). The last portion of the of *speB* CDS (156 nt) was not removed in order to keep the RBS of the following gene (*SPy_2038*) intact. PCR-mediated ligation with OLEC7565/OLEC7569 was used to ligate the upstream and downstream regions with the lox71-*ermAM/B*-lox66 cassette, which was amplified from the vector pEC454 using OLEC1899/OLEC1900. The fragment obtained was cloned in BamHI (Fermentas)-digested pSEVA141, which is a suicide vector for *S. pyogenes* (*i.e.* it does not replicate in this bacterium) (see section above for cloning strategy). Prior transformation in *S. pyogenes*, the obtained plasmid was linearized by cleaving in the *amp*R cassette with SacII (New England Biolabs). Insertion of the lox71-*ermAM/B*-lox66 cassette was checked by PCR and DNA sequencing (Microsynth, Switzerland). Elecro-compentent *S. pyogenes* cells, from an erythromycin resistant clone, were transformed with pEC455 expressing the Cre recombinase, in order to remove the *ermAM/B* cassette resulting in the lox72 site. Kanamycin-resistant and erythromycin-sensitive clones were grown without antibiotic at 37°C to lose the pEC455 vector. The integrity of the generated strain was checked by both PCR and DNA sequencing (Microsynth, Switzerland).

2.2.2 Construction of plasmids for the *speB* promoter study

The plasmid-based reporter system (pLZ12Km2-P23R:TA:*ffluc*) described in a previous study[357] was used as a backbone to generate plasmid pEC2248, in which the *ffluc* (firefly luciferase gene) expression is controlled by the *speB* promoter region (P_{speB}). This region, containing both *speB* promoters (P and P1), comprises the DNA portion from − 940 to − 697 nt relative to *speB* start codon and contains the predicted binding sites of the transcriptional regulator RopB[301,315]. pLZ12Km2-P23R:TA:*ffluc* was digested with SacI and SacII (Thermo Scientific) to eliminate the lactococcal promoter P23, a constitutive promoter upstream of the *ffluc* ORF[356,357]. P_{speB}, inclusive of P and P1, was amplified from WT genomic DNA using primers containing the SacI and SacII restriction sites (OLEC8386/OLEC8387) and cloned (see section 2.2) in the digested pLZ12Km2-P23R:TA:*ffluc*.

2.2.3 Construction of plasmids for *speB* ectopic expression

The shuttle vector pEC85 was used to ectopically express *speB* in the Δ*speB* strain[328,353]. *speB* (including 5' UTR and CDS) was amplified from WT genomic DNA using primers OLEC7970/7971. The fragment was ligated by PCR-mediated ligation with OLEC7968/OLEC7971 to the *gyrA* promoter (P_{gyrA}) from *Streptococcus agalactiae*, which was amplified from pEC455 using OLEC7968/7969. The resulting fragment was cloned (see section 2.2) in XbaI and EcoRI (Fermentas)-digested pEC85, to generate plasmid pEC2146, which was used as a template to construct the variant plasmids with mutations in *speB* 5' UTR. These variant plasmids were generated according to the two-stage PCR mutagenesis protocol[358] with the Pfu Turbo DNA polymerase (Stratagene) and primers indicated in Table 6. 25 µl from the PCR reaction were then treated with 10U of DpnI (Fermentas) for at least 1h at 37°C and 5 µl were transformed in *E. coli* chemically competent cells as descried in section 2.1.2.

2.3 RNA techniques

2.3.1 RNA extraction

The *S. pyogenes* WT and deletion mutant overnight cultures were diluted 1:200 in 300 ml of THY and harvested at early-logarithmic (EL) ($OD_{620\ nm}$= 0.1), mid-logarithmic (ML) ($OD_{620\ nm}$ = 0.25) and early-stationary (ES) ($OD_{620\ nm}$= 0.4) phases of growth (see Figure 7A). In particular, 25 ml of cultures were mixed with 25 ml of 1:1 ice-cold acetone (Sigma)/ethanol (AppliChem) solution, centrifuged at 3500 g for 10 min at 4°C and washed twice with 5 ml and 1 ml of ice-cold

TE-sucrose buffer (50 mM Tris-HCl pH 8.0, 10 mM EDTA, pH 8.0, 50 mM NaCl, 25% sucrose), respectively. The cell lysis was performed incubating for 5 min on ice and at 95°C for 1.5 min with 100 µl of lysis buffer (20 mM Tris-HCl pH 8.0, 50 mM EDTA pH 8.0, 20% sucrose, 5 µl of lysozyme (Sigma) 50 mg/ml, 5 µl of mutanolysin (Sigma) 10 µg/µ) and 150 µl of lysis executioner (2% Dodecyl sulfate sodium salt (SDS; Applichem) and 1 mg/ml Proteinase K (Fisher Scientific)), respectively. The lysed cells were mixed with 750 µl of TRIzol (Life Technologies) and incubated at room temperature for 5 min. Subsequently, 200 µl of chloroform (Sigma) were added and mixed with the samples by vertexing for 15 sec. The mixture was incubated at room temperature for 10 min and centrifuged at 11300 g for 10 min at 4°C. Following the centrifugation, the sample separated into three phases: upper aqueous phase containing the RNAs, an interphase and lower organic phase. The upper aqueous phase (approximately 700 µl) was transferred into a fresh tube and the RNAs were precipitated by adding 700 µl of ice-cold isopropanol (Sigma-Aldrich) and incubating at least 1 h at − 20°C. The RNAs were pelleted by centrifugation conducted at 11300 g for 10 min at 4°C. The supernatant was removed without disturbing the RNA pellet and subsequently washed with 1 ml of 70% EtOH and centrifuged at 15800 g for 5 min at 4°C. The supernatant was discarded and the RNA pellet was air-dried at room temperature for at least 10 min and resuspended in approximately 50 µl of autoclaved Milli-Q H_20. The RNA concentrations were measured using a NanoDrop (Thermo Scientific) and the RNA integrity was analyzed by agarose gel (2%) electrophoresis conducted at 100 V for 30 min and staining with ethidium bromide (Merck).

2.3.2 Rifampicin assay

Overnight cultures were diluted 1:200 in 250 ml of medium and the bacteria were grown until mid-logarithmic and/or early-stationary growth phases. At these time points, the rifampicin (Sigma-Aldrich), which was previously freshly dissolved in methanol (Sigma-Aldrich), was added to a final concentration of 250 µg/ml. 25 ml of cultures were harvested immediately (0 min) and at specific time points (5, 10, 20, 30, 45 min or 1, 2, 4, 8 min) after the addition of rifampicin, by mixing with 1:1 ice-cold acetone (Sigma)/ethanol (Applichem) solution and centrifuging at 3500 g for 10 min at 4°C. Total RNA was extracted as described in the previous section 2.3.1 and used to perform Northern blot analyses.

2.3.3 Agarose Northern blot analysis

The RNAs (15 µg or 20 µg) were mixed to 2X RNA loading buffer, (95% Formamide (Sigma Aldrich), 0.025% SDS (AppliChem), 0.025% Bromophenol Blue (Merck) and 0.5 mM

EDTA) and were separated on an agarose gel (1% agarose, 30 ml of 10X MOPS (200 mM 3-N-morpholinopropanesulfonic acid, 50 mM sodium acetate and 10 mM EDTA, pH 7), 6.6% formaldehyde), in 1X MOPS running buffer (20 mM 3-N-morpholinopropanesulfonic acid, 5 mM sodium acetate and 1 mM EDTA, pH 7) supplemented with 0.7% formaldehyde (Merck). The gel was pre-run for 15 min at 80 V and the running procedure was conducted at the same voltage for 2h. The gel was then washed in autoclaved water (3X for 5 min, at room temperature) and incubated in 20X SSC (3 M NaCl and 0.3 M sodium citrate pH 7.0) for 5 min at room temperature. The RNAs were transferred onto a Nylon Hybond N+ membrane (GE Healthcare) using a capillarity system consisting of a large container, filled with 1L of transfer buffer (20X SSC) and multiple layers of absorbing material. In particular, three sheets of Whatman paper (Omnilab) were pre-wet with 20X SSC and placed on top of a support positioned within the large container, in order to allow the contact of the Whatman papers with the transfer buffer. After the incubation in 20X SSC, the gel was placed on top of the three Whatman papers. The Nylon Hybond N+ membrane was pre-wet in both autoclaved MQ-H_20 and 20X SSC and subsequently positioned on top of the gel. Additional Whatman papers (3X), of the size of the gel, were pre-wet in 20X SSC and placed on top of the Nylon Hybond N+ membrane. In all the above-mentioned steps, the air bubbles between the different layers of the system were removed, to allow homogeneous transfer of the RNAs to the membrane. Finally towel papers and weights were placed on top of the absorbing material. This set-up was used to transfer the RNA overnight at room temperature. The RNAs were UV-crosslinked to the 6X SSC-rinsed membrane using the UV Stratalinker 1800 (Stratagene) with 2X autocrosslinking mode. The oligonucleotides probes (Table 7) were radioactively labelled as was previously described[359]. In brief, 2 µl oligonucleotides (from a 20 pmol/µl solution) were first denatured (10 min at 95°C and on ice, respectively) and then incubated with, 2 µl of gamma-^{32}P ATP (Hartmann analytics), 1 µl of T4 Polynucleotide Kinase (T4-PNK, Fermentas) (10 U/µl) and 2 µl of T4 PNK-Buffer A (10X) at 37°C for 30 min in a final volume of 20 µl. The labelled oligonucleotides were purified over G-25 columns (GE Healthcare), following the manufacturer's instructions, in order to remove the unincorporated nucleotides. Following incubation at 95°C for 10 min and subsequently on ice for other 10 min, the probes were incubated with the membranes in approximately 15 ml of hybridization buffer (Rapid-hyb buffer, GE Healthcare). The hybridization was performed overnight at 50°C with rotation in a hybridization oven (Analytic Jena). The membranes were then rinsed twice with pre-warmed 1X SSC-0.1% SDS buffer and washed for 20 min at 50°C with pre-warmed 1X SSC-0.1% SDS buffer and then with pre-warmed 0.5X SSC-0.1% SDS for 20 min. Visualization of the radioactive signal was done after approximately one week of exposure to a BSA phosphor screen (GE Healthcare) using a

phosphorimager (Typhoon Fla 9000, Fujifilm). 16S rRNA was used as a loading control. The approximate sizes of the RNAs were estimated using the RiboRuler High Range Ladder (Thermo Scientific), which was radioactively labelled with 5' ^{32}P pCp (Hartmann analytics) using T4 RNA Ligase I (NEB) overnight at 16°C. The radioactively labelled ladder was then purified using the RNeasY MinElute Kit (QIAGEN) following the manufacturer's instructions. Each Northern blot was performed at least in biological triplicates.

Oligo	Sequence 5′-3′ [a]	F/R [b]	Usage [c]	Target [d]
Northern blot analyses				
OLEC5802	*AACCACATAGTAGGCGCCTC*	R	NB	*speB* 5′ UTR
OLEC7431	*GCAACACATCCTGTAGCTGC*	R	NB	*speB* CDS
OLEC6215	*TGGTTATCATACGATGTCGC*	R	NB	*speB* 5′ UTR
OLEC2410	*ATTTGCTTTTAACGGTACAT*	R	NB	*speB* 5′ UTR
OLEC5803	*GGTGCATAAGGTCAATAGCCA*	R	NB	*ropB* 5′ UTR
OLEC1542	*CATGACACGATTCATATTAGTC*	R	NB	*covR* CDS
OLEC8864	*CTAAGCGACTACCTTATCTCA*	R	NB	TPP riboswitch
OLEC8857	*TCACAGCATCCACAGACTCT*	R	NB	*serS*
OLEC8860	*CGATTTCCGCGGTACCAC*	R	NB	*SPy_1570-SPy_1569-valS*
OLEC8856	*CTTTCCACCAACCAGCACTC*	R	NB	*thrS* 5′ UTR
OLEC8866	*TTAATCACACTGCCCAAACGG*	R	NB	*glyQ* 5′ UTR
OLEC5805	*TATCTTCGATACCGCCCAAG*	R	NB	*rpsB* CDS
OLEC8344	*CTTCGACTTGCAGGCTAACC*	R	NB	*rpsB* 5′ UTR
OLEC5807	*TCGTTGTTAGCTGGTTTGCC*	R	NB	*tsf* CDS
OLEC5799	*ACAGCCACTGACGCTAAACC*	R	NB	*bmpA* CDS
OLEC5797	*TCTGCCATTACCTGACGACA*	R	NB	*cdd* CDS
OLEC10381	*TTGCGCTGCCTTCAATGAAT*	R	NB	*SPy_1551* CDS
OLEC10384	*GATAGTGATGCCCGCTTGTT*	R	NB	*murC* CDS
OLEC10501	*TGCCACCGTTCTTACCGTA*	R	NB	*SPy_0316* CDS
OLEC8749	*CACGTTGAATATTGCCAGCTTCA*	R	NB	*pyrH* CDS
OLEC8754	*CGACTGTCAGTGCTATTTCGCC*	R	NB	*SPy_2197*
OliRN243	*CGTTGTACCAACCATTGTAGC*	R	NB	16S rRNA
OLEC288	*CTAAGCGACTACCTTATCTCA*	R	NB	5S rRNA

Table 7. **Oligonucleotides used as probes in Northern blot analyses.**
The name, the sequence, the directionality, the function and the respective targets of the oligonucleotide probes used in this thesis are reported. [a] *italic*: sequence annealing to the template; [b] F: forward primer; R: reverse primer; [d] 5' UTR: 5' untranslated region; CDS: coding DNA sequence.

2.3.4 Polyacrylamide Northern blot analysis

10 µg of total RNAs mixed 2X RNA loading buffer were separated on 10% or 8% polyacrylamide – 8 M urea gels in Tris-Borate-EDTA (TBE) 1X buffer (89 mM Tris Base, 89 mM Boric Acid, and 2 mM EDTA) for approximately 1h and 30 min at 100 V. RNAs were then transferred on Nylon Hybond N+ membrane (GE Healthcare) either by electro-blotting (Trans-Blot SD semi-dry transfer apparatus, Bio-Rad) using Whatman papers pre-wet with TBE for 1h at 18 V or by wet transfer using a Biorad Trans-Blot Cell system in 1X TBE for 1h and 15 min at 50 V. The RNA molecules were chemically cross-linked to the membranes at 60°C for 1h using

EDC (N-(3-Dimethylaminopropyl)-N-Ethylcarbodiimide Hydrochloride, Sigma) (0.375 g per membrane) dissolved in H_20 (pH adjusted to 8.0) and 12.5 M 1-Methylimidazole (Sigma). Alternatively, the RNAs were cross-linked as described above (section 2.3.3) using UV. Oligonucleotides probes (Table 7) were labelled with ^{32}P using the T4 Polynucleotide Kinase (T4-PNK, Fermentas) as described above (see section 2.3.3). After pre-incubating the membranes with approximately 15 ml of hybridization buffer (Rapid-hyb buffer, GE Healthcare) at 42°C for at least 1 h, the labelled oligonucleotides were added. The hybridization was carried out overnight at 42°C, with rotation in a hybridization oven (Analytic Jena). The membranes were then washed for 15 min at 42°C with washing buffer I (5X SSC, 0.1% SDS) and washing buffer II (1X SSC, 0.1% SDS) for additional 15 min at 42°C. The BSA phosphor screen (GE Healthcare) was exposed for at least one week and the signal was visualized with a phosphorimager (Typhoon Fla-9000, Fujifilm). The 5S rRNA served as a loading control. The RNA approximate sizes were estimated according to the 30-330-bp AFLP® DNA ladder (Invitrogen) or to the RNA decade™ Marker (Invitrogen). Both ladders were radioactively labelled according to the manufacturer's instructions, respectively.

2.3.5 RNA half-life measurements

The transcript abundance was measured by densitometry with FiJi[360]. The obtained values, deriving from three independent blots, were fitted on an exponential decay curve and the half-life was calculated. The *speB* transcript abundance comparison between the WT and Δ*rny* strains was performed using a *t-test*.

2.3.6 Primer extension

10 µg of total RNAs were denatured with subsequent incubations at 95°C and on ice for 10 min, respectively. The denatured RNAs were then annealed with 3 µl oligonucleotides (Table 8), previously radiolabeled as described in section 2.3.3., in the presence of 0.8 mM dNTPs, at 65°C for 30 min and chilled for 1 min on ice. The reverse transcription was performed using 200 U of Superscript III Reverse Transcriptase (Invitrogen), 1X first strand buffer, 5 mM Dithiothreitol (DTT) and 40 U of RNaseOUT Recombinant Ribonuclease Inhibitor (Invitrogen), for 1h at 55°C in a final volume of 20 µl. The cDNA products were precipitated with 0.1 volume of 3M NaAc pH 5.2 (room temperature) and 3 volumes of 100% ethanol (ice-cold) and subsequently centrifuged at 20000 g for 30 min at 4°C. After removal of the supernatant, the pellets were washed with 1 ml of 70% ethanol and centrifuged at 20000 g for 10 min at 4°C. The supernant was discarded and the pellets were air-dried for 10 min and resuspended in 10 µl of 2X RNA loading dye (95% formamide, 0.025% SDS, 0.025% bromophenol blue and 0.5 mM EDTA). The samples were

denatured at 95°C for 5 min and resolved on a 10% polyacrylamide – 8 M urea gels in TBE 1X buffer. The cDNA approximate sizes were estimated using 30-330-bp AFLP® DNA ladder (Invitrogen). Alternately, sequencing ladder was generated with a dideoxy chain termination reaction, using Sequenase™ Version 2.0 DNA Sequencing Kit (Affymetrix USB), according to the manufacturer's instructions. The DNA template used during the termination reaction was generated by PCR from WT genomic DNA with OLEC3903/OLEC3970 (Table 8).

Oligo	Sequence 5′-3′ [a]	F/R [b]	Usage [c]	Target [d]
Primer extension analyses				
OLEC3966	*TGTCATATGTTAAACCTTTCTAATC*	R	PE	*ropB* 5′ UTR
OLEC2406	*ACTACCATTTTGCAAAAGGAAC*	R	PE	*speB* 5′ UTR
OLEC3903	*TAACGGTACATTGGACACACCTCC*	R	PE	
OLEC3904	*TATACCTCTTTCAAAATAAGTTAATCTACTGC*	R	PE	
OLEC3970	*TGGGTTAGCAAGAACAAATCC*	R	PE	*speB* CDS
OLEC4030	*GTGTTTGTCTTTTGGCGATAACC*	R	PE	*SPy_0116* CDS
OLEC4031	*CACTGACGCTAAACCAAGAC*	R	PE	*bmpA* CDS
OLEC4033	*TCCTAGCTGTCCATCTGTGC*	R	PE	SPy_sRNA1680670
OLEC4036	*AATGCTTCATTCTTGAGTTGGC*	R	PE	*SPy_1845* CDS

Table 8. **Oligonucleotide primers used in primer extension analyses.**
The name, the sequence, the directionality, the function and the respective targets of the oligonucleotides used in primer extension analyses are reported. [a] *italic*: sequence annealing to the template; [b] F: forward primer; R: reverse primer; [c] PE: primer extension analysis; [d] 5' UTR: 5' untranslated region; CDS: coding DNA sequence.

2.3.7 Simultaneous mapping of RNA 5' and 3' ends

Total RNA was extracted as described above (see section 2.3.1) and treated with Turbo DNase (4 U/ µl) (TURBO DNA-*free*™, Ambion) according to the manufacturer's instructions and the RNA integrity was checked using a bioanalyzer (Agilent). 1 µg of DNase-treated RNA was incubated with 1X Thermopol buffer (NEB) in presence or absence of 50 U of RNA 5' pyrophosphohydrolase (RppH; NEB), which transforms the 5' triphosphate (5' PPP) of primary transcripts to 5' monophosphate (5' P), in a final volume of 20 µl at 37°C for 1h. Autoclaved MQ-H_20 was added to the samples to reach a final volume of 200 µl and the RNAs were purified using phenol-chloroform purification. 200 µl of Acid Phenol:Chloroform (5:1; Ambion) were mixed with the samples by vertexing. The mixture was subsequently centrifuged at 10270 g for 10 min at 4°C and divided in an aqueous phase on top and an organic phase at the bottom. The aqueous phase (approximately 200 µl) was transferred to a fresh reaction tube, mixed with 200 µl of Chloroform (Sigma) and centrifuged at 10270 g for 10 min at 4°C. Repeatedly, the aqueous phase was transferred in a fresh reaction tube, mixed with 0.1 volume of 3M NaAc pH 5.2 (at room temperature), 3 volumes of 100% ethanol (ice-cold), 0.5 µg/µl Glycogen (Thermo Fisher) and incubated for at least 1 h, at − 20°C. The samples were then centrifuged at 10270 g for 30 min at

4°C and the supernatant were discarded. After a wash step with 1 ml of ice-cold 80% EtOH and centrifugation at 10270 g for 5 min at 4°C, the RNA pellet was air-dried at room temperature for 10 min and resuspend in 12 µl of autoclaved MQ-H_2O and the RNA concentration was determined using a NanoDrop (Thermo Scientific). 50 ng of RppH-treated and non-treated RNA were self-ligated at 17°C overnight with 40 U T4 RNA Ligase (NEB), 1X Ligase buffer, 8% v/v DMSO, 10 U Ribonuclease inhibitor (RNaseOUT, Thermo Fisher Scientific) and 1 U DNase I (NEB), in a final volume of 20 µl. Following phenol-chloroform extraction and ethanol precipitation, the self-ligated RNAs were retrotranscribed and the cDNAs amplified using OneStep RT-PCR kit (QIAGEN) following the manufacturer's instructions with specific pairs of outward and inward (used as a control) primers listed in Table 9. RT-PCR products were visualized on a 2% TAE (0.4 M Tris acetate, pH 8.3, and 0.01 M EDTA) agarose gel and 5 µl of the RT-PCR reaction was used to clone the RT-PCR products in the vector of the TOPO TA cloning kit (Invitrogen) according to the manufacturer's instructions. Up to ten clones were selected and analyzed by DNA sequencing (Microsynth, Barcode Economy Run Service) with two different primers (OLEC3224/ OLEC3225) listed in Table 6. To map the RNA 5' and 3' ends, the sequences were compared with *S. pyogenes* M1 SF370 genome.

Oligo	Sequence 5′-3′ [a]	F/R [b]	Usage [c]	Target [d]
Simultaneous mapping of RNA 5′ and 3′ ends				
OLEC7573 (inward)	*CATACGATGTCGCTGATACCATG*	R	Circ-RT-PCR	*speB* 5′ UTR
OLEC7574 (inward)	*ACCGTTAAAAGCAAATGCAGTAG*	F	Circ-RT-PCR	
OLEC7575 (outward)	*ATGATAACCATACGATTCAGCTAAG*	F	Circ-RT-PCR	
OLEC7576 (outword)	*TTTTAACGGTACATTGGACACAC*	R	Circ-RT-PCR	

Table 9. **Oligonucleotide primers used in the simultaneous mapping of RNA 5' and 3' ends.**
The name, the sequence, the directionality, the function and the respective targets of the oligonucleotides used in the simultaneous mapping of the 5' and 3' ends are reported. [a] *italic*: sequence annealing to the template; [b] F: forward primer; R: reverse primer; [c] Circ-RT-PCR: circularization of RNA by self-ligation followed by Reverse Transcriptase-Polymerase Chain Reaction; [d] 5' UTR: 5' untranslated region.

2.4 Protein detection

2.4.1 Luciferase reporter assay

The plasmids constructed to study the activity of *speB* promoters (described above) were transformed in *S. pyogenes* WT and Δ*rny* competent cells as described above (see section 2.1.2). Single clones, containing the plasmid, were subsequently grown in THY supplemented with kanamycin. WT and Δ*rny* cells derived from a single clone and containing the constructs for studying the activity of *speB* promoters were grown overnight in 2 ml of THY with kanamycin. 1 ml was used to inoculate 50 ml medium in 100 ml-flasks for the over-day cultures. At mid-

logarithmic and early-stationary growth phases, beetle luciferin (Promega) was added, to a final concentration of 50 ng/µl, to 200 µl of cultures in a white opaque 96-well microtiter plate (Greiner Bio-One™). Luminescence was measured, at room temperature, using a microplate reader (BioTek™ Cytation 3) with the following conditions: integration time of 1 sec, gain of 120 and read height of 1 mm. The signal was normalized with the $OD_{620\ nm}$ and with the luminescence measured from bacterial cells with the pEC2173 construct harboring the constitutive promoter P23. The experiments were performed in biological triplicates, each with technical triplicates and the standard deviation was calculated.

2.4.2 Exoprotein precipitation

20 ml of bacterial cultures were harvested at mid-logarithmic and early-stationary growth phases, centrifuged at 3500 g for 10 min at 4°C and the respective supernatants were filtered with a 0.45 µm syringe filter (VWR). The proteins, contained in the cell-free supernatants, were precipitated with a final concentration of 10% of ice-cold trichloroacetic acid (Sigma) for 30 min on ice. After centrifugation at 8400 g for 10 min at 4°C, the proteins were washed with ice-cold acetone, air-dried and resuspended in 70 µl of 1 M Tris-HCl, pH 8.0. Equal volumes of protein preparation, derived from the different *S. pyogenes* cultures, were mixed with 5X SDS loading dye (250 mM Tris-HCl pH 6.8, 10% SDS, 0.02% of bromophenol blue, 30% (v/v) glycerol, 5% β-mercaptoethanol (Sigma)), incubated at 95°C for 5 min and subsequently on ice. The proteins were resolved on a 15% polyacrylamide gel by sodium dodecyl sulfate – polyacrylamide gel electrophoresis (SDS-PAGE). SDS-PAGE gels are composed of a separation gel at the bottom and a stacking gel on top. The separation gel solution was prepared as follows: 2.1 µl 1M Tris-HCl at pH 8.8, 1.875 ml of 40% polyacrylamide solution, 50 µl of 10% SDS, 0.975 ml MQ-H_2O, 50 µl of 10% ammonium persulfate solution (APS; AppliChem) and 3 µl of N,N,N',N'-tetramethylethylene-diamine (TEMED; Sigma-Aldrich). When the separation gel was completely polymerized, the stacking gel solution was prepared (390 µl of 1M Tris-HCl at pH 6.8, 450 µl of 40% polyacrylamide solution (AppliChem), 30 µl of 10% SDS, 2.13 ml MQ-H_2O, 30 µl of 10% APS and 1.8 µl TEMED) and added on top of the separation gel. The protein run was conducted at 80 V for 25 min in the stacking gel and at 100 V for 2.5 h, in 1X SDS running buffer (25 mM Tris-HCl, 250 mM glycine, 0.1% SDS). The proteins in the separation gel were stained with Coomassie Brilliant Blue (Sigma-Aldrich), following the manufacturer's instructions, and imaged with a scan (CanoScan LiDE 700F). The PageRuler™ Plus Prestained Protein ladder (Thermo Fisher) was used to approximately estimate the protein masses in kilodaltons (kDa).

2.5 RNA sequencing

2.5.1 Library preparation

Total RNA was prepared, as described above (see section 2.3.1), from the WT and Δ*rny* strains grown until mid-logarithmic growth phase and from WT grown until early-stationary phase growth and depleted from Genomic DNA using Turbo DNase (Ambion), following the manufacturer's instructions. Following sample quality control using a bioanalyzer (RNA 6000 Nano Kit, Agilent) and Qubit (RNA BR Assay, Invitrogen) the rRNAs were removed with Ribo-Zero rRNA Removal Kit (Illumina). 5' PPP RNAs were converted into 5' P RNAs using either TAP (Tobacco Acid Pyrophosphatase, Epicentre) (for the WT sample at early-stationary growth phase) or using RppH (New England Biolabs) (for the WT and Δ*rny* samples at mid-logarithmic growth phase). The TAP treatment was conducted in a microcentrifuge tube and in a 50 µl final volume, including 25 µl of total RNA derived from the previous step, 5 µl of 10X TAP reaction buffer, 1 µl of TAP enzyme (10 U) and 19 µl nuclease-free water. The RppH was similarly conducted in a 50 µl final volume, including 25 µl of total RNA, 5 µl of 10X RppH reaction buffer, 2 µl of RppH enzyme (10 U) and 18 µl of nuclease-free water. Both the mixtures were incubated at 37°C for 1.5 h and the RNAs were then purified using Phenol/Chloroform extraction as follows. 250 µl of nuclease-free water was added to the reaction and mixed with 300 µl of Phenol:Chloroform:Isoamylalcohol (25:24:1; Roth), at pH 4.5-5, by vertexing for 10 sec. The samples were then centrifuged at 15800 g for 10 min at room temperature. The obtained aqueous phase, was transferred in a fresh reaction tube and mixed with 300 µl of Chloroform:Isoamylalcohol (24:1; Sigma) by vertexing for 30 sec. The samples were then centrifuged at 15800 g for 5 min at room temperature and the aqueous phase transferred in a fresh reaction tube. The last two described steps (washes with Chloroform:Isoamylalcohol and centrifugation) were repeated for a second time. The RNAs were then precipitated by mixing the aqueous phase with 1/10 volume of 3M NaAc, 2.5 µl Glycogen (20 mg/ml; Fermentas), 3 volumes of ice-cold EtOH and subsequently incubating the samples at – 80°C for 1 h or overnight at – 20°C. Following RNA precipitation, the samples were centrifuged at 15800 g for 30 min at 4°C. The respective supernatants were discarded and 500 µl of ice-cold 70% EtOH was added to the pellets. After centrifugation at 15800 g for 5 min, the supernatants were removed and the pellets were air-dried from 5 to 15 min. The TAP-treated and the RppH-treated RNAs were resuspended in 25 µl of TE-buffer and in 15.2 µl of nuclease-free water, respectively. 15 µl from each of the RNA samples were treated with T4 Polynucleotide kinase (T4 PNK; Thermo Scientific) in order to phosphorylate the RNA 5' ends and to remove the 3' phosphoryl groups. The reaction was carried out for 25 min at 37°C in presence of 2 µl 10X buffer A, 1.8 µl 10mM ATP, 1 µl T4 PNK.

The RNAs were then purified using the RNA Clean & Concentrator kit (Zymo Research), following the manufacturer's instructions and separated in two different fractions (short and long RNAs). The fraction of long RNAs was fragmented either by chemical fragmentation using NEB Next Magnesium RNA Fragmentation Module (NEB E61505) (for the WT sample at early-stationary growth phase) or by mechanical fragmentation for 140 s using Covaris M220 and microtube AFA Fiber Pre-Slit Sbak-Cap tubes (for the WT and Δ*rny* samples at mid-logarithmic growth phase). The cDNA libraries were prepared using NEXTflex® Small RNA Sequencing Kit v3 (Bioo Scientific), The manufacturer's instructions were followed until step G, consisting of 22 cycles of PCR. After replacing step H with the final purification of the cDNA library, executed using Agencourt AMPure XP beads (Beckman Coulter), instructions of step I from the protocol of the NEXTflex® Rapid Directional qRNA-Seq Kit (Bioo Scientific) were followed. The cDNA library quality control was assessed using Qubit (dsDNA HS Assay, Invitrogen) and a bioanalyzer (High sensitivity DNA kit, Agilent). The preparation of the cDNA libraries was performed by K.H. For all samples the high-throughput RNA sequencing was executed on a HiSeq3000 with a paired-end mode and 150 nt of read-length (for the WT sample at mid-logarithmic growth phase) and 75 nt of read-length (for the WT and Δ*rny* samples at early-stationary growth phase) at the Max Planck-Genome-centre Cologne. The data have been deposited at NCBI under the accession number SRP149897 and SRP149896 for the WT sample at early-stationary growth phase and for the WT and Δ*rny* samples at mid-logarithmic growth phase, respectively.

2.5.2 Read processing

FastQC (v0.11.5) was used to analyze the quality of the data. Reads were filtered and eliminated if the quality score was <10 and if smaller than 22 nt. The sequences corresponding to the adapters were removed using Cutadapt (v1.11)[361]. The mapping of the reads was performed with STAR (v2.5.2b) in both 'random best' and 'end to end' modes[362]. NC_002737.2 was used as *S. pyogenes* reference genome[363]. Samtools (v1.3.1) was used to sort and index the obtained BAM files. Including four random bases at the RNA 3' and 5' ends, during the library preparation, allowed the detection and filtering of potential PCR artefacts. These bases serve as Unique Molecular Identifiers UMIs and they were identified applying the UMI Tools (v0.4.1). The coverage files (total, 5′ and 3′ ends) were generated with the HTSeq library (v0.9.1) using a custom script and were visualized with the Integrative Genomics Viewer (IGV)[364]. The read processing procedure was performed by R.A.-B. from the MPUSP bioinformatic research platform.

2.5.3 Differential expression analysis

Gene counts were determined using FeatureCounts (v1.5.2)[365] with the *S. pyogenes* gene annotation NC_002737.2 in which the sRNAs, previously identified[359] were included. Genes that were differentially expressed in the Δ*rny* strain compared to the WT strain were identified utilizing edgeR (v3.20.6)[366,367] using an absolute log2 fold change (log2 FC) ≥ 1 and a false discovery rate (FDR) < 0.05. The differential expression analysis was performed by R.A.-B, from the MPUSP bioinformatic research platform, and analyzed by A.L.R., A.-L.L. and L.B.

2.5.4 Identification of RNase Y processing sites

The identification of processing positions was performed, using a method developed in our laboratory[185,204]. After data normalization, the differential expression analysis of the 5' and 3' read ends was performed by comparing the data deriving from triplicate samples of the WT and Δ*rny* strains (WT *vs* Δ*rny* and Δ*rny vs* WT). When a 5' or 3' RNA end was identified as more abundant in the WT strain than in Δ*rny* strain it was annotated as 5' or 3' *rny*_end, respectively (Tables A2 and A4). Conversely when a 5' or 3' RNA end was more abundant in the Δ*rny* strain than in the WT strain it was mapped as 5' or 3' Δ*rny*_end, respectively (Table A10). A "counts per million (cpm)" value ≥ 0.05 was set for each comparison in order to prefilter the genome coverage. Only the 5' and 3' end positions that have a cpm ≥ 5 were further analyzed. The differential expressed ends were obtained using edgeR (v3.20.6) with an absolute log2 fold change (FC) ≥ 1 and false discovery rate (FDR) < 0.05. The obtained results were filtered by applying previously developed parameters to identify specific cleavage positions[185] including: (i) the proportion of ends ≥ 2% (reads stopping at the position of interest compared to all the reads mapping at that position is least 2%), (ii) the ratio of proportion of ends between WT and Δ*rny* ≥ 3 (ensuring that the position of interest is identified as a cleavage position independently of the RNA abundance).

2.5.5 Comparison of RNase Y and 3'-to-5' exoRNases targetomes

The identified 5' and 3' *rny*_ends were compared to the PNPase, YhaM and RNase R targetomes (*i.e.* 3'-to-5' exoRNase trimming start and stop positions), which were previously identified by our laboratory[204] (see Figure 4) and the RNA sequencing data relative to the 3'-to-5' exoRNases are deposited in NCBI under the accession number (SRP149887). Different strategies were used to compare these RNA ends. When at least two successive positions were identified as 3' *rny*_ends in a window of 5 nt, the position with the highest ratio of proportion of ends between the WT and Δ*rny* strains was selected for following analyses. First, the 3' *rny*_ends were compared with the PNPase, YhaM and RNase R trimming stop positions, allowing for a shift of +/− 5 nt

(see Table A5). Second, the 3' *rny*_ends matching 3'-to-5' exoRNase trimming stop positions, identified above, were compared with the trimming start positions located downstream (see Table A6). A 200 nt maximum distance between these trimming stop and start positions was set for PNPase and RNase R, and 10 nt for YhaM. In the case of RNase R, by allowing a maximum distance of 200 nt, no trimming start positions downstream of the 3' *rny*_ends, which matched RNase R stop positions, were identified. Third, the 3' *rny*_ends were compared with the PNPase, YhaM and RNase R trimming start positions, allowing for a shift of +/− 5 nt (see Table A7). Finally, the 5' *rny*_ends were paired to the PNPase, YhaM and RNase R trimming start positions found 10 nt upstream (see Table A9). The comparisons, described above, were performed with Python (v3.6.3) by T.T.R together with A.L.R., A.-L.L. and L.B.

2.5.6 Sequence alignments and RNA secondary structure prediction

The sequence logos were generated using WebLogolib (v3.5.0)[368] with sequences of 20 nt centred at the processing site and a GC content of 38.5%. The minimum fee energy (MFE, ΔG in Kcal mol^{-1}) was calculated using a sliding window of 50 nt sequences, 100 nt upstream and downstream of the processing position of interest, using RNAfold (v2.4.3)[369]. The average MFE at each nt was then calculated. The sequence and structural analysis of the sequences of interest was performed by T.T.R together with L.B. A.L.R. and A.-L.L

The structure predication of the *speB* mRNA 5' UTR was conducted using RNAfold (v2.4.3)[369] and by calculating the ΔG in Kcal mol^{-1} of 25 nt sequences 100 nt upstream and 100 nt downstream of the RNase Y processing position at – 137 nt from the *speB* mRNA start codon. The RNA structure of a portion of *speB* mRNA 5' UTR including the sequence from – 153 nt to the *speB* start codon was predicted with RNAfold WebServer and visualized using VARNA[370].

Results

3.1 Characterizing the global functions of *S. pyogenes* RNase Y

3.1.1 Deletion of *rny* affects the expression of virulence genes

To study the role of RNase Y in *S. pyogenes*, the gene encoding this endoRNase (*SPy_1633*, named here *rny*) was deleted, generating the Δ*rny* strain (constructed by A.-L.L.). The *rny* gene is not essential for the growth of *S. pyogenes*, as also reported for *S. aureus*[167] and *B. subtilis*[162] and unlike *C. perfringens* where even the depletion of the enzyme causes severe effects on growth[166]. We observed that the Δ*rny* strain was delayed in growth compared to the wild type (WT) strain under the conditions tested (Figure 7A).

To examine the global impact of RNase Y on the transcriptome, differential gene expression analysis was performed. Due to the difference in growth between the Δ*rny* and WT strains, the cultures for the RNA isolation were harvested when the same optical density ($OD_{620\,nm}$ 0.25), corresponding to mid-logarithmic phase of growth, was reached regardless of the growth time (Figure 7A and B). The differential gene expression analysis revealed that 80 genes that were significant differentially expressed, with an absolute log2 fold change (log2 FC) ≥ 1 and a false discovery rate (FDR) < 0.05 (Table A1). Among the differentially expressed genes, 22 and 58 genes were down and upregulated, respectively. In the absence of RNase Y, the expression of genes involved in different physiological processes, including nucleotide and vitamins metabolism, were altered. For instance, the *folEPQK* operon, involved in the folate synthesis, was upregulated in the Δ*rny* strain compared to the WT strain (Table A1). In *S. pyogenes*, the *folEPQK* transcript was previously shown to be stabilized in the Δ*rny* strain[371]. In addition, genes involved in protein synthesis and folding were also upregulated in the Δ*rny* strain, including genes encoding ribosomal proteins (*i.e. rpsB* and *rpsL*), the elongation factor thermos-stable EF-Ts (*tsf*) and protein chaperons (*groEL* and *groES*) (Table A1).

We did not identify any major transcriptional regulator (see also section 1.3.7) whose mRNA abundance was altered in the Δ*rny* strain (Table A1). However, we observed that the expression of 6 different putative sRNAs was affected in the absence of RNase Y (Table A1).

Of note, the expression of genes encoding for *S. pyogenes* virulence factors (*slo*, *speB*, *grab* and *sic*) was affected in the Δ*rny* strain. In particular, *speB* and the neighboring genes in the chromosome (*SPy_2040*, *SPy_2041* and Spy_sRNA1699993) were strongly downregulated in the Δ*rny* strain (Figure 7B and Table A1).

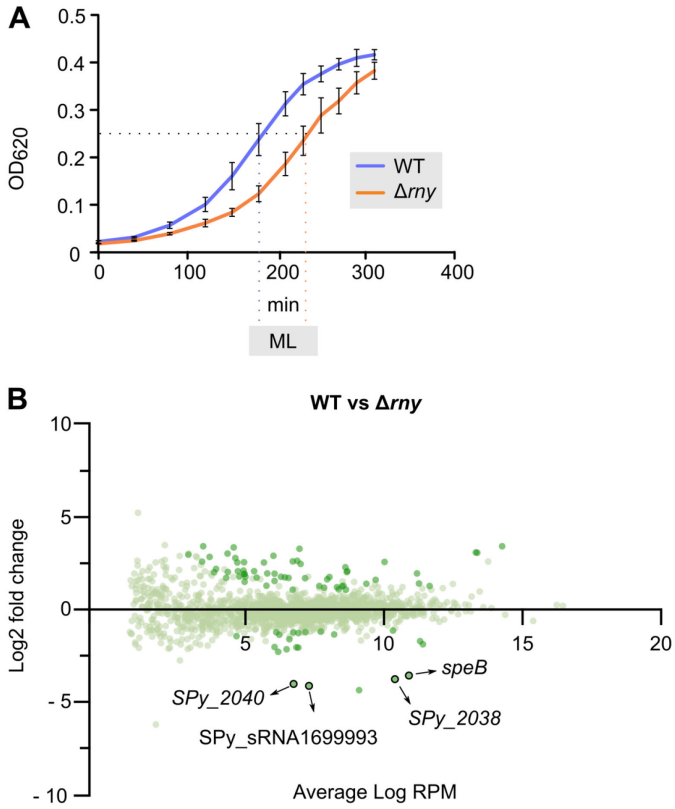


Figure 7. **Effects of *rny* deletion in *S. pyogenes*.**
A. Growth curve analysis of the wild type (WT) and *rny* (RNase Y) deletion mutant (Δ*rny*) strains in THY at 37°C with 5% CO_2 in triplicates. The error bars indicated the standard deviation of the mean. At mid-logarithmic (ML) phase growth the samples were harvested to perform RNA sequencing. K.H. contributed in the replicate generation of the *S. pyogenes* growth curves. **B.** Mean difference plot of the WT strain *versus* the Δ*rny* strain, showing the log-fold change and the average log reads per million reads (RPM), a unit that reflects transcript abundance for each gene (represented as dots). The dark green dots represent the genes that were further selected with an absolute log2 fold change ≥ 1 and false discovery rate < 0.05 and that were also identified in the comparison Δ*rny versus* complemented *rny* deletion strain (Δ*rny*::*rny*). *speB* and the genes expressed within the same polycistronic transcript (*SPy_2040*, SPy_sRNA1699993 and *SPy_2038* were downregulated in the Δ*rny* strain compared to the WT strain. Differential expression analysis was performed by R.A.-B and analysed by A.L.R, A.-L.L. and L.B.

3.1.2 RNase Y regulates *speB* expression

To give insights into the role of RNase Y in *speB* regulation, the *speB* expression pattern during the bacterial growth, was evaluated by Northern blot analyses and the production of the SpeB protease was analysed by SDS-PAGE (Figure 8). As mentioned in the introduction, it is well known that *speB* expression is growth phase-dependent and tightly controlled by a plethora of factors and environmental conditions[290]. Consistent with the observation that *speB* expression depends on the bacterial growth phase, *speB* mRNA abundance increased from early-logarithmic to early-stationary growth phase (Figure 8).

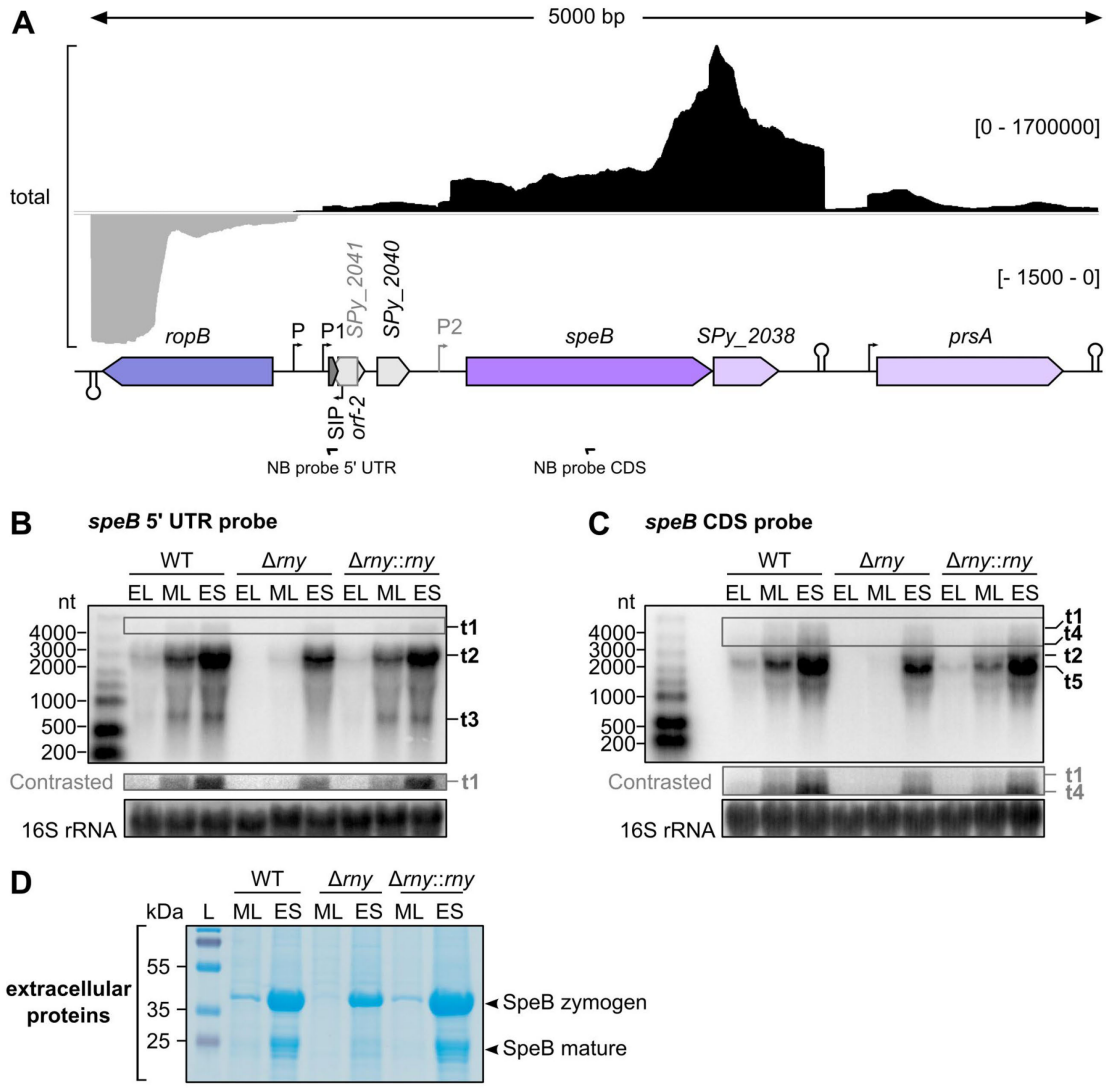


Figure 8. **RNase Y controls *speB* expression.**
A. Total coverage profiles of *speB* and surrounding genes resulting from RNA sequencing (black for the positive strand and grey for the negative strand). The coverage scale is indicated between brackets. The genes (arrows) with the predicted promoters and terminators are shown. *ropB* (Regulator of protease B) is a transcriptional activator of *speB* expression[301]; *SPy_2038* was shown to encode for an inhibitor of SpeB protease activity; *prsA* encodes for the PrsA foldase, which is involved in the maturation process of SpeB

protease. The *ropB-speB* intergenic regions harbours several ORFs: *SPy_2040*, *orf-2*, the SpeB inducing peptide (SIP) and *SPy_2041*. P and P1 are the *speB* transcriptional start sites[161] and P2 depicts an RNA 5′ end, which was previously identified as a *speB* TSS[301], but then annotated as a putative cleavage site[161]. The arrows below the locus indicate the labelled oligonucleotide probes used in the Northern blot analyses (NB probe) and binding either to the *speB* 5′ untranslated region (5′ UTR) or to the coding DNA sequence (CDS). **B** and **C.** *speB* expression profile and transcript abundance was investigated by Northern blot analyses performed in the WT, *rny* (RNase Y) deletion mutant (Δ*rny*) and *rny* chromosomal complemented deletion mutant (Δ*rny*::*rny*) strains that were grown until early-logarithmic (EL), mid-logarithmic (ML) and early-stationary (ES) growth phases. The different *speB* transcript isoforms are indicated (from t1 to t5). The poorly detectable isoforms are shown in a contrasted portion of the blot. The 16S rRNA was used as a loading control. The results of one representative Northern blot analysis (n>3) are shown. **D.** The effect of RNase Y deletion on SpeB protease levels was studied by analysing the extracellular fraction on a sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) stained with Coomassie blue in the WT, Δ*rny* and Δ*rny*::*rny* strains. The SpeB zymogen (SpeBz, 40 kDa) and the SpeB mature form (SpeB, 28 kDa) are indicated on the right side of the gel. The results of one representative SDS-PAGE gel (n=3) are shown. K.H. contributed in the extracellular protein extraction and SDS-PAGE analysis.

To facilitate the identification of the *speB* transcript isoforms (*i.e.* transcript variants originating from the same locus), Northern blot analyses were performed using a labelled oligonucleotide probe binding to either the *speB* mRNA 5' UTR or to the *speB* CDS, respectively (Figure 8A, B and C).

First, by using the probe binding to the *speB* mRNA 5' UTR, ~ 4kb transcript isoforms (t1), deriving from the co-transcription of *speB* with *SPy_2038* and *prsA2*, were observed. Transcription driven from *speB* promoters until the predicted terminator downstream of the *SPy_2038*, led to the production of the t2 transcript isoforms (Figure 8B and C). Smaller isoforms (t3) were also visible and it is likely that the 5' end of these transcripts coincides with the previously annotated P2 transcriptional start site (TSS), but recently annotated as a putative cleavage site[161]. Two additional isoforms (t4 and t5) were detected when the probe binding to the *speB* CDS was used, but their origin is unclear (Figure 8C). The abundance of *speB* transcript was reduced at mid-logarithmic and early-stationary growth phases in the Δ*rny* strain when compared to the WT and to the Δ*rny*::*rny* strains, confirming the RNA sequencing data (Figure 8B and C). In agreement with the *speB* transcript levels, the amount of both the zymogene and the mature forms of extracellular SpeB protease was decreased in the Δ*rny* strain and was restored in the Δ*rny*::*rny* strain (Figure 8D).

3.1.3 RNase Y controls *speB* promoter activity

The downregulation of *speB* expression observed in the Δ*rny* strain (Figure 8) could be caused by (i) an indirect effect of RNase Y on *speB* expression (*e.g.* RNase Y affects a regulator of *speB*), (ii) a direct effect (*e.g.* RNase Y processes the *speB* transcript) and (iii) both scenarios. To examine whether the reduction in *speB* abundance in the Δ*rny* strain was due to an indirect transcriptional effect, the luciferase reporter gene (*ffluc*) was fused to the *speB* promoters and *ffluc*

gene expression was measured in the WT, Δ*rny* and Δ*rny*::*rny* strains (Figure 9A and B). The luminescence signal was strongly reduced in the Δ*rny* strain when compared to the WT and the Δ*rny*::*rny* strains, during both mid-logarithmic and early-stationary growth-phases (Figure 9B). This result indicated that *speB* promoters were less active in the Δ*rny* strain than in the WT strain. The luminescence signal was instead equal in the WT and Δ*speB* strains, indicating that putative SpeB-mediated autoregulatory circuits were not occurring at the experimental conditions used. (Figure 9C). All together these results demonstrated that RNase Y is indirectly involved in promoting the activation of *speB* promoters.

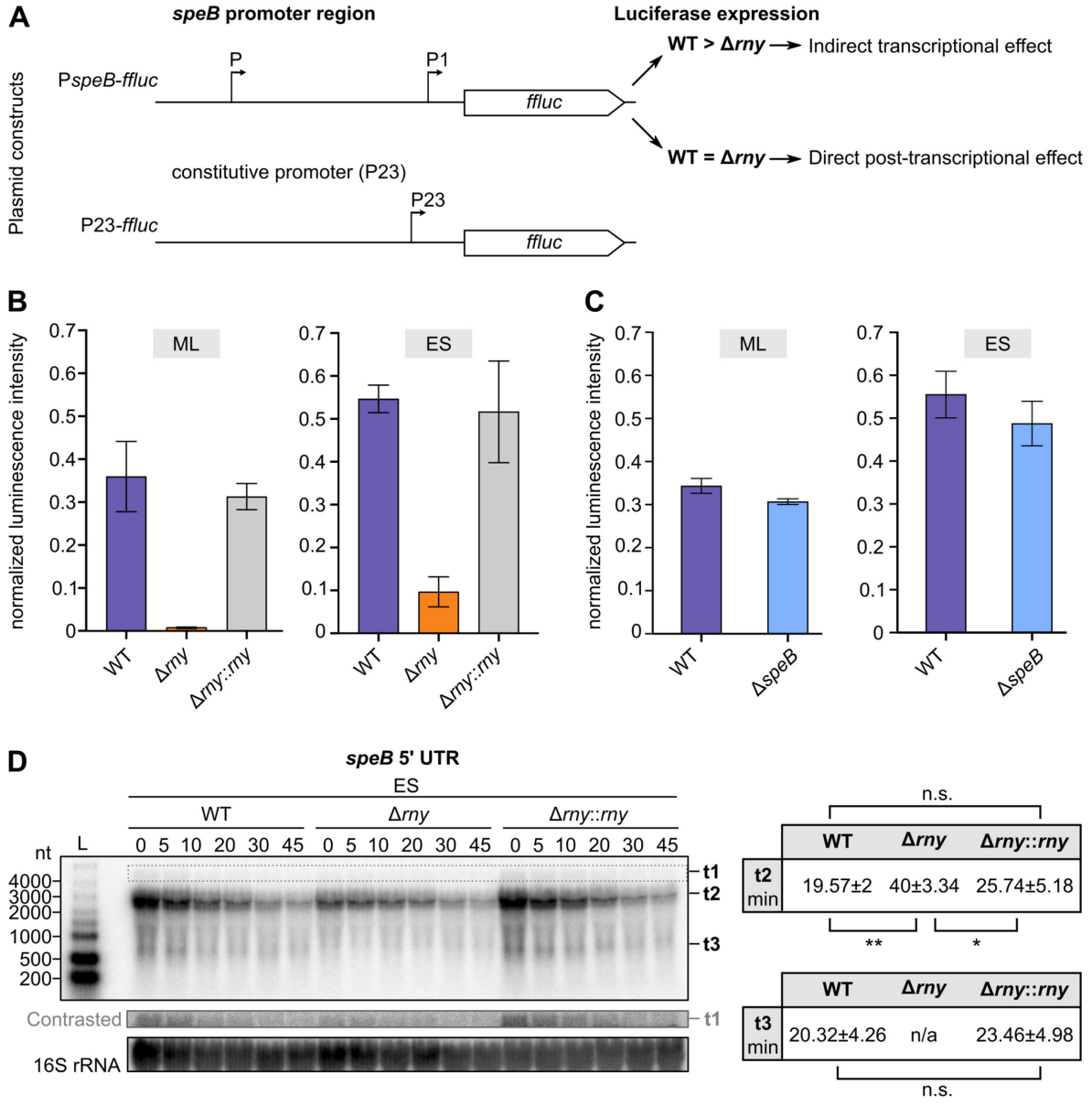


***Figure 9.* RNase Y regulates *speB* expression indirectly at the transcriptional level.**
A. Schematic representation of the luciferase (*ffluc*) fusion vectors used to study the *speB* promoter activity in the WT and *rny* (RNase Y) deletion mutant (Δ*rny*) strains. The *speB* promoter region (containing P and P1 TSSs) were cloned upstream of the *ffluc*. The cloned region also comprises the putative binding sites of

the RopB transcriptional activator[301]. A control vector (P23-*ffluc*) in which the *ffluc* expression is under the control of a constitutive promoter (P23) was used in the experiment. **B** and **C.** Luminescence assay to study *speB* promoter activity in the WT, *rny* (RNase Y) deletion mutant (Δ*rny*) and *speB* (SpeB) deletion mutant (Δ*speB*) strains at both mid-logarithmic (ML) and early-stationary (ES) phases of growth. The luminescence intensity of the samples is relative to the plasmid control (P23-*ffluc*) and normalized to the $OD_{620\ nm}$. The experiment was performed in three biological replicates each with technical triplicates and the standard deviation is shown. S.M., master student in the Charpentier group, contributed in the luminescence assay experiments. **D.** Analysis of *speB* transcript stability using rifampicin assay at ES growth phase in the WT and Δ*rny* strains. The minutes after the addition of the rifampicin are indicated. The region of *speB* mRNA 5' UTR targeted by the labelled oligonucleotide probe in the Northern blot analysis is shown in Figure 8A. The *speB* transcript isoforms are indicated (see also Figure 8B) and the 16S rRNA was used as a loading control. The results of one representative Northern blot analysis (n=3) are shown. The calculated half-life of the t2 and t3 isoforms are reported in the right panel. The half-life of the t1 isoform could not be calculated due to the low intensity of the band and the high background of the blot. The values of the half-life, determined from three independent experiments, were compared using a *t-test*. Single and double asterisks depicted a p-value inferior of 0.05 and 0.01, respectively. The n.s. indicates that there is not significant difference between the samples.

Next, to decipher whether the reduction in *speB* abundance in the Δ*rny* strain was also due to a direct role of RNase Y on *speB* mRNA, the *speB* transcript stability was analysed by rifampicin assay followed by Northern blot analyses at early-stationary growth phase using the labelled oligonucleotide probe binding to the *speB* mRNA 5' UTR (Figure 9D). The *speB* transcript isoforms were not destabilized in the absence of RNase Y, on the contrary the half-life of the *speB* isoforms t2 was increased in the Δ*rny* strain (40 min ± 3.34) compared to the WT strain (19.57 min ± 2) (Figure 9D). Taken together, these results indicate that in *S. pyogenes* SF370 strain, RNase Y causes *speB* downregulation mainly indirectly at the transcriptional level. However, the observation that the stability of *speB* isoform t2 increased in the Δ*rny* strain suggested that RNase Y might also affect *speB* expression at the post-transcriptional level. Interestingly, when *speB* expression pattern was examined by Northern blot analysis with a labelled oligonucleotide probe binding to the *speB* mRNA 5' UTR, the smaller isoforms t3 were not detectable in the Δ*rny* strain (Figure 9D), suggesting that RNase Y was involved in their production and that they likely resulted from the processing of the transcript isoforms t2. This hypothesis would explain the stabilization of the t2 isoforms in the Δ*rny* strain and will be corroborated in the next sections.

3.1.4 The expression of some *speB* regulators is affected by *rny* deletion

RNase Y affects the *covRS* transcript stability

Several transcriptional regulators were described to control *speB* expression (Table 1 and Figure 10A) and the effect of RNase Y on few of these regulators has been described so far. A microarray analysis of RNA samples extracted from cultures treated with rifampicin was previously performed to estimate the half-life of *S. pyogenes* RNAs in both the WT and Δ*rny* strains[106]. The transcripts encoding for the CovRS two-component system, composed of the response regulator

(CovR) and sensor kinase (CovS) were reported to be stabilized, however this result was never validated[106]. CovRS was shown to repress the *speB* transcription by directly targeting the *speB* promoter region[307,308,335]. Here, the stability of the *covRS* transcript was analysed by Northern blot analysis at mid-logarithmic growth phase, using a labelled oligonucleotide probe binding the *covR* CDS (Figure 10B). One long (tl) and one short (ts) transcript isoforms were identified by Northern blot analysis (Figure 10B). The stability of the long isoform was equal in the WT and Δ*rny* strains. This transcript could result from the co-transcription of *covR* and *covS*, which were already demonstrated to be transcribed as a sole mRNA[372]. Indeed, a terminator for *covR* and a promoter for *covS* were not identified in the short intergenic region between the two genes. The short transcript isoform, containing the *covR* CDS, was highly increased in the Δ*rny* strin and the transcript was still detectable up to 45 min after the addition of the rifampicin, compared to the WT strain (Figure 10B). This result suggested that the *covR* transcript is possibly an RNase Y direct target and that the CovRS system could be involved in the RNase Y-mediated regulation of *speB* expression.

RNase Y and RNase III process the *ropB* mRNA

speB expression can be activated by the stand-alone transcriptional regulator RopB, which is encoded adjacent to *speB*. Based on the observation that *ropB* overexpression in the Δ*rny* strain re-established the SpeB production to the WT level, RopB was proposed to be involved in the RNase Y control of *speB* expression[168]. Analysis of *ropB* transcript stability revealed that *ropB* mRNA was highly stabilized in the Δ*rny* strain (Figure 10C, left), as already reported in *S. pyogenes*[106,161]. Whereas in the WT strain only one *ropB* transcript isoform (t2) was detected, two additional transcript isoforms (t1 and t3) were present and highly stabilized in the Δ*rny* strain (Figure 10C, left). Based on the increase of *ropB* transcript stability in the Δ*rny* strain, it is likely that RNase Y is the endoRNase responsible for the initiation of *ropB* transcript degradation.

Next, we evaluated the *ropB* expression pattern in the WT strain, in three different growth phases and we observed a reduction of *ropB* transcript abundance from early-logarithmic to early-stationary growth phases (Figure 10C, right). Conversely, the analysis of the *ropB* expression profile in the Δ*rny* strain unveiled that the *ropB* abundance increased during growth, indicating that RNase Y controls the *ropB* temporal expression pattern (Figure 10C, right).

The *ropB* mRNA 5' UTR was suggested to be targeted by endoRNases, including RNase Y[106,106,161]. We mapped the RNA 5' ends in the *ropB* mRNA 5' UTR by RNA sequencing and primer extension analyses, at mid-logarithmic growth phase (Figure 10D). The *ropB* TSS was annotated at – 369 nt from the *ropB* start codon and the signal intensity of the cDNA product

corresponding to this position was increased in the Δ*rny* strain compared to the WT, Δ*rny*::*rny* and the *rnc* (RNase III) deletion mutant (Δ*rnc*) strains (Figure 10D). This observation could be explained by the higher level of *ropB* transcript in the Δ*rny* strain during mid-logarithmic phase of growth (Figure 10C, right). Importantly, a low-abundance RNA 5' end, at − 218 nt from *ropB* start codon, was observed by primer extension in the WT, Δ*rny*::*rny* and Δ*rnc* strains, but not in the Δ*rny* strain, indicating that this 5' end resulted from RNase Y processing of the *ropB* mRNA 5' UTR (Figure 10D, see contrasted portion), in accordance with a previous study[106]. A second RNA 5' end was not present in the Δ*rny* strain, indicating that it was also likely produced by RNase Y (Figure 10D, see contrasted portion).

Additional RNA 5' ends were observed in the *ropB* mRNA 5' UTR and probably corresponding to other RNase processing positions. One of these RNA 5' ends was not detected in the Δ*rnc* strain indicating that was likely generated by RNase III. The processing by RNase III, which is a dsRNA-specific endoRNase, generates two new RNA 5' ends. While one RNA 5' end was detected, the second was not retrieved by primer extension analysis (Figure 10D). It is possible that the second RNase III processing position was not detected because: (i) RNase III cleaves the RNA duplex formed by *ropB* mRNA 5' UTR and another RNA molecule, (ii) RNase III nicks only one strand of a stem loop structure, (iii) the second RNA 5' end is rapidly degraded, or (iv) the primer used in the primer extension analysis did not allow to capture this RNA 5' end. The *ropB* mRNA 5' UTR is predicted to be highly structured[106] and the RNase Y and RNase III processing events could be responsible for a structural rearrangement of this region, which could have consequences, for instance, on the *ropB* transcript translation and/or degradation.

In summary, RNase Y likely regulated *ropB* expression and transcript stability through the processing of the *ropB* mRNA 5' UTR (Figure 10C and D). If the higher *ropB* mRNA stability in the Δ*rny* strain causes an increase in the RopB protein levels, then RopB, which is an activator of *speB* transcription, would not be involved in the RNase Y-mediated regulation of *speB* expression. However, it must be noted that in the Northern blot analysis, performed after the rifampicin assay, three different *ropB* transcript isoforms were identified and while two were highly stabilized in the Δ*rny* strain, one was destabilized in this strain (Figure 10C). It is unknown whether the three isoforms are equally translatable or if only one specific isoform is efficiently translated, therefore the RopB protein level cannot be deduced from the sole increase of *ropB* transcript abundance and stability.

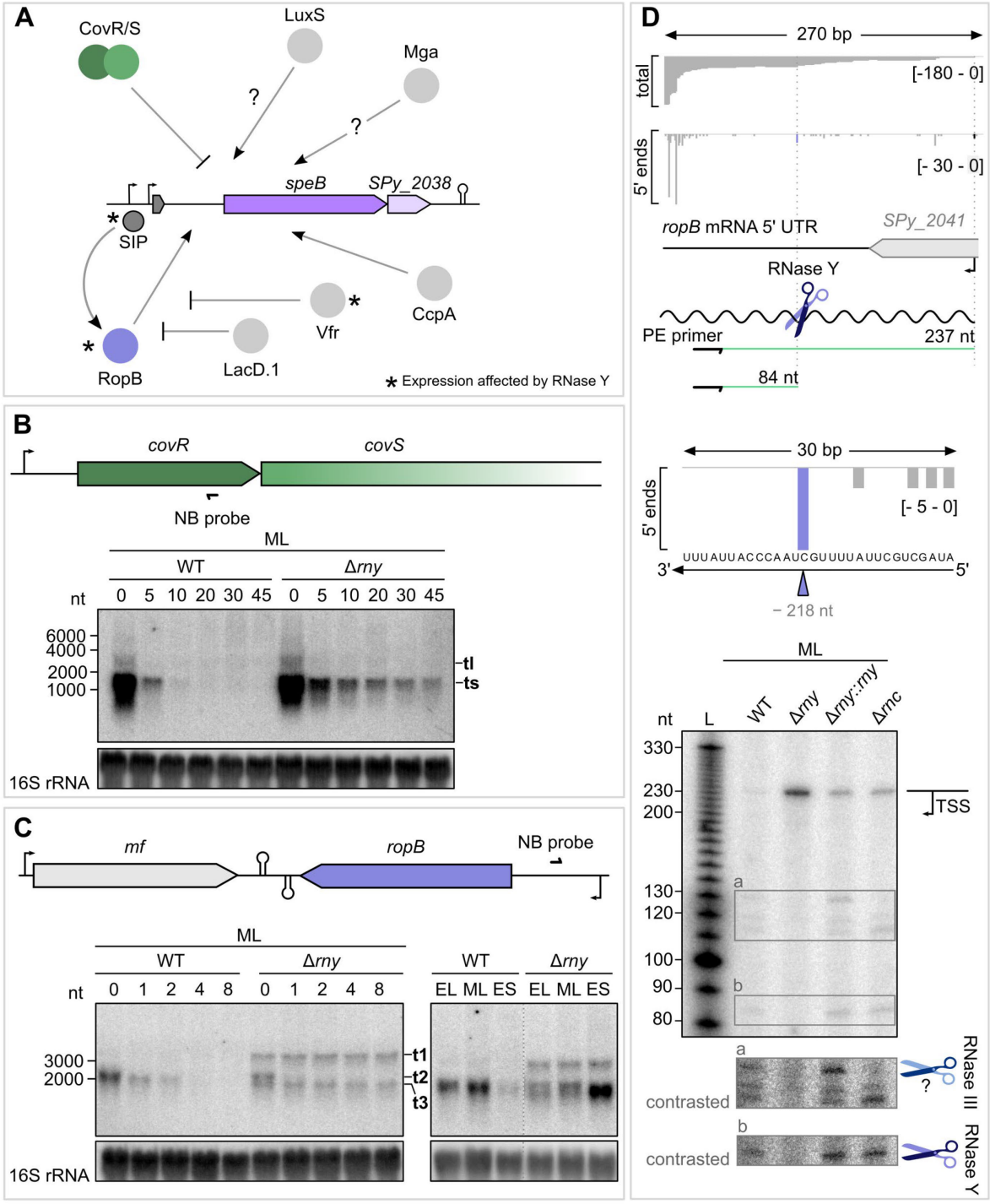


Figure 10. **The mRNA stability of the RopB and CovR regulators is affected in the Δ*rny* strain.** **A.** Schematic representation of the major *speB* transcriptional regulators (see section 1.3.6 and Table 1). The arrows indicate whether the regulator activates (black arrowhead) or represses (no arrowhead) *speB* expression. RopB, CovR and CcpA are direct transcriptional regulators, acting at the *speB* promoter region. Vfr, LacD.1 and SIP are indirect *speB* regulators, acting through RopB. The question marks indicate that the mechanism of *speB* regulation is not characterized yet. The asterisk depicts the *speB* regulators whose expression is affected by RNase Y. **B.** and **C.** Study of *ropB* (in panel B) and *covR* (in panel C) transcript stability in the WT and *rny* (RNase Y) deletion mutant (Δ*rny*) strains at mid-logarithmic (ML) phase of growth. The pattern of *ropB* expression was analysed at early-logarithmic (EL), ML and early-stationary (ES) growth phases. The minutes after addition of the rifampicin are indicated. The labelled oligonucleotide probes used in the Northern blot analyses (NB probe) are show below the schematic representation of the

loci. The 16S rRNA was used as a loading control. The results of one representative Northern blot analysis (n=3) are shown. The study of *covR* mRNA stability was performed by L.B. and A.-L.L. **D.** Analysis of *ropB* mRNA 5' UTR by RNA sequencing (top and bottom). Total and 5' end coverage profiles from RNA sequencing of a region of *ropB* mRNA 5' UTR which includes the putative *SPy_2041* ORF. The black bent arrow depicts the *ropB* transcriptional start sites (TSS) at – 369 nt from *ropB* start codon. The coverage scales are indicated between brackets. A zoom of the RNA 5' end coverage profile showing the putative RNase Y cleavage site in *ropB* mRNA 5' UTR at 218 nt upstream of the *ropB* start codon is shown below. The expected cDNA products obtained in the primer extension analysis and the primer used in the experiment are represented with green lines and a black arrow, respectively. The primer extension analysis was performed in the WT, Δ*rny*, Δ*rny::rny* and *rnc* (RNase III) deletion mutant (Δ*rnc*) strains at ML growth phase. The results of one representative primer extension analysis (n=3) are shown. RNase Y and RNase III putative cleavage sites are indicated by purple and blue scissors, respectively. For the poorly detectable cDNA products of interest, a contrasted portion is shown.

3.1.5 Conclusions I

In the absence of RNase Y, 80 genes were differentially expressed, including *speB*, which was strongly downregulated in the Δ*rny* strain. A transcriptional reporter system analysis revealed that RNase Y regulates *speB* expression mainly indirectly at the transcriptional level through the activity of a so far unidentified intermediate factor. The expression of two major *speB* regulators (RopB and CovRS) is controlled by RNase Y and we hypothesized that CovRS could be involved in the RNase Y-mediated regulation of *speB* expression, without excluding the possibility that other factors are also involved in this process. Finally, from the analysis of *speB* transcript isoform stability, we deduced that RNase Y likely also cleaves the *speB* mRNA.

3.2 Identifying the requirements for RNase Y processing

3.2.1 Mapping of RNA 5' ends in the WT and Δ*rny* strains: RNase Y preferentially cleaves after a guanosine

The differential expression analysis performed in the WT and Δ*rny* strains revealed genes which are regulated by RNase Y in a direct and/or an indirect manner (Figure 7B and Table A1). We next aimed to specifically identify, at a single nucleotide resolution, the direct targets of RNase Y (*e.g.* RNAs which are processed by this endoRNase). To pinpoint the RNase Y processing positions genome-wide, the abundance of the RNA 5' ends in the WT, Δ*rny* and Δ*rny*::*rny* strains was compared, in three biological replicates at mid-logarithmic phase of growth, using a method previously developed in our laboratory[185,204] (Figure 11A and B, see Material & Methods). The RNA 5' ends that were more abundant in both the WT and Δ*rny*::*rny* strains than in Δ*rny* strain were annotated as 5' *rny*_ends (Figure 11A and B). When the RNA abundance was too low either in the WT strain or in Δ*rny* strain, the processing positions were not annotated (see Material & Methods). We identified 190 RNA 5' ends whose abundance was RNase Y-dependent and therefore likely resulted from RNase Y processing (Table A2). Selected 5' *rny*_ends were validated by primer extension analysis (Figure 11C), which was performed in the WT and Δ*rny* strains at mid-logarithmic growth phase. The expected cDNA products, from the primer to the predicted RNA 5' end, were observed for the four tested targets (Figure 11C).

The RNA 5' end generated by RNase Y after processing of the *rpoB* mRNA 5' UTR previously identified by primer extension (Figure 10D) was not identified here (Table A2). This can be explained by the observation that the *rpoB* mRNA 5' UTR was poorly expressed (see coverage profile in Figure 8A and 11D) and as already mentioned, positions located in regions not sufficiently expressed were excluded from the analysis (see Material & Methods)[185,204].

To evaluate whether RNase Y exhibited any structural or sequence preference, we analysed the sequences surrounding the 5' *rny*_ends. To examine the presence of putative secondary structures around the 5' *rny*_ends, the minimum free energy (ΔG in Kcal mol^{-1}) was estimated 100 nt upstream and downstream of the 5' *rny*_ends (Figure 11D, see Material & Methods). An overall increase of the ΔG was observed at the 5' *rny*_ends compared to the neighboring regions, indicating that the RNase Y processing occurred at ssRNA regions (Figure 11D). Sequence alignment of the 190 5' *rny*_ends revealed the preference of a guanosine one nucleotide upstream of the cleavage site in 87.4 % of the cases. This result indicates that this nucleotide likely plays an important role in RNase Y processing of the target RNA (Figure 11D).

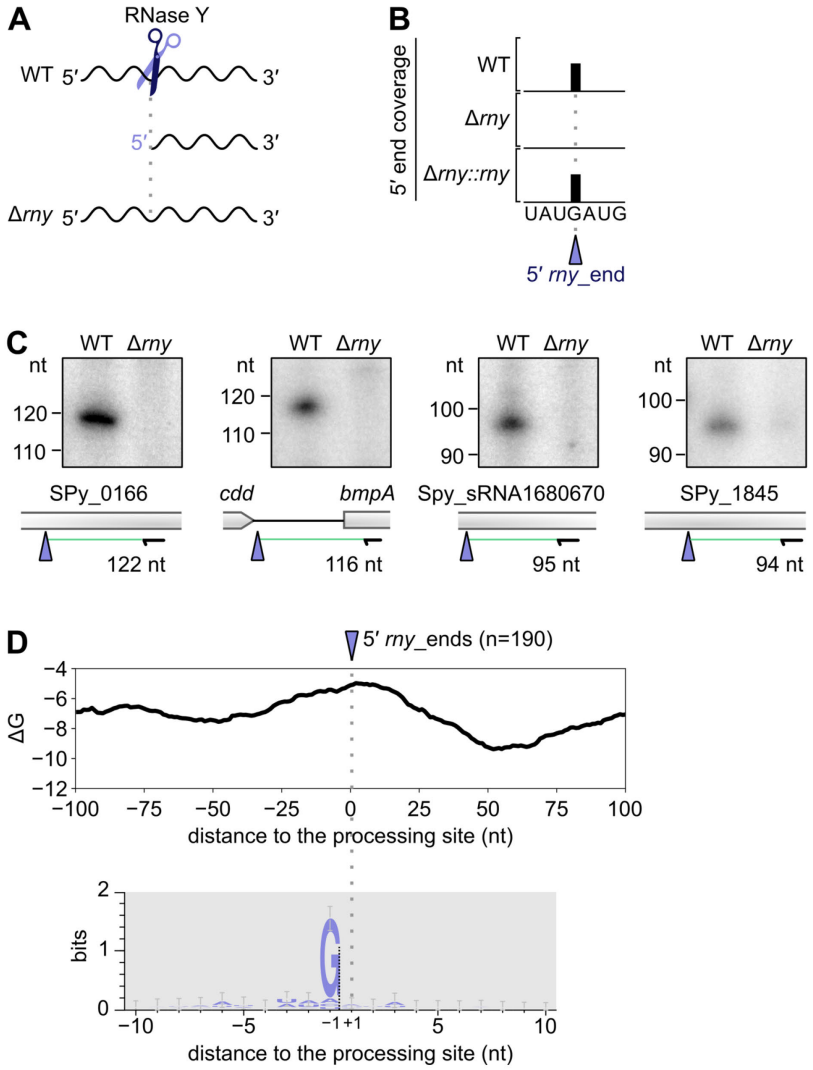


Figure 11. **RNase Y preferentially cleaves after a guanosine.**
A. Example of an RNA molecule which is processed by RNase Y (scissors) generating a new RNA 5' end detectable in the WT strain but not in the *rny* (RNase Y) deletion mutant (Δ*rny*) strain. **B.** Representation of RNA sequencing 5' end profile. The RNA 5' end more abundant in the WT and in the complemented *rny* deletion mutant (Δ*rny::rny*) strains than in the Δ*rny* strain were annotated as 5' *rny*_end (purple arrowhead). The "UAUGAUG" represents a random sequence cleaved by RNase Y. **C.** Subset of 5' *rny*_ends that were selected and validated by primer extension analysis performed at mid-logarithmic growth phase in the WT and Δ*rny* strains. The primer used is indicated in black and the green lines depict the length of the cDNA product from the primer to the predicted 5' *rny*_end. The results of one representative primer extension analysis (n=3) are shown. The validation of the 5' *rny*_end by primer extension analysis was performed together with A.-L.L. Note that for all the targets tested other RNA 5' ends (sometimes coinciding with the TSS) were detected upstream of the RNase Y processing positions by primer extension analysis in both the WT and Δ*rny* strains, confirming that the RNA analysed were expressed in both strains. **D.** Structure (on the top) and sequence (on the bottom) conservation of the 5' *rny*_ends. The minimum free energy (ΔG, in kcal/mol^{-1}) was calculated at each nucleotide using a sliding window of 50 nt, 100 nt upstream and downstream of the 5' *rny*_ends. The logo was generated by aligning 10 nt each side of the 190 5' *rny*_ends. The sequence and structure analyses at the 5' *rny*_ends were conducted by T.T.R. with A.L.R., A.-L.L. and L.B.

3.2.2 Mapping of the RNA 5' ends in *speB* mRNA 5' UTR

The study of *speB* expression pattern performed in section 3.1.2 revealed that the 5' end of the t3 transcript isoforms might derive from RNase Y processing of the *speB* primary transcript (t2) (Figure 8A-C and 9D). In the genome-wide mapping of RNase Y-dependent RNA 5' ends, two RNA 5' ends generated by RNase Y were identified in the *speB* CDS (Table A2). However, these positions did not correspond to the 5' termini of the t3 transcript isoforms and therefore the origin of these transcripts remained unclear. To elucidate the mechanism of the *speB* transcript isoform generation, the *speB* 5' UTR was re-annotated by mapping the *speB* mRNA 5' boundaries at early-stationary growth phase, during which *speB* is highly expressed.

First, the *speB* mRNA 5' ends coinciding with the *speB* TSSs were annotated (Figure 12). Two *speB* TSSs, located at 697 and 842 nt from the *speB* start codon, were identified by RNA sequencing and confirmed by primer extension analyses at early-logarithmic, mid-logarithmic and early-stationary growth phases (Figure 12A and B). At early-stationary phase of growth additional RNA 5' ends were also detected by primer extension analysis between the positions corresponding to the predicted TSSs P and P1 (Figure 12A and B), thus it is possible that they correspond to processing sites or to alternative TSSs.

The – 10 boxes for P and P1 promoters and a putative − 35 motif for the P promoter were manually predicted (Figure 12C). A previous study of the *speB* promoter region identified a − 35 motif for the P1 promoter, which was located at 16 nt upstream of the – 10 box[301]. In our strain, a canonical − 35 motif upstream of the P1 promoter was absent (Figure 12C). RopB binding sites were identified upstream of the P and P1 promoters[301,315], indicating that RopB can initiate *speB* expression from both the annotated promoters (Figure 12C). These results are in agreement with previous studies, since P1 (− 697 nt) was already described as *speB* TSS[301] and position P (− 842 nt) was recently captured by 5' rapid amplification of cDNA ends[161].

The mapping of the *speB* TSS (Figure 12A) and *ropB* TSS (Figure 10D) confirmed that the two promoters overlap and that these two genes are divergently transcribed (Figure 8A). The long *ropB-speB* intergenic region (940 nt) encompasses several small ORFs. Originally, the two putative ORFs *SPy_2040* and *SPy_2041* were annotated upstream of *speB* and *ropB*, respectively[363] (Figure 8A and Figure 12C). However, a more recent study proposed a new annotation with *SPy_2040* (named *orf-3*), SpeB inducing peptide (SIP) and *orf-2*, but not *SPy_2041*, which was not reported[315]. The RNA sequencing data and the primer extension analysis with the primer targeting the RNA downstream of the *ropB* TSS did not reveal a TSS for *SPy_2041*. In addition, the *ropB* TSS was located downstream of the *SPy_2041* start codon, therefore this ORF could not be co-transcribed

with *ropB* (Figure 10D). Beside SIP, which is involved in the RopB-mediated regulation of *speB* expression, the function of the other putative ORFs is currently unknown.

In conclusion, we re-mapped the features of the *S. pyogenes* SF370 serotype M1 *speB-ropB* intergenic region, which has been not coherently annotated[106,161,301,303,315,331].

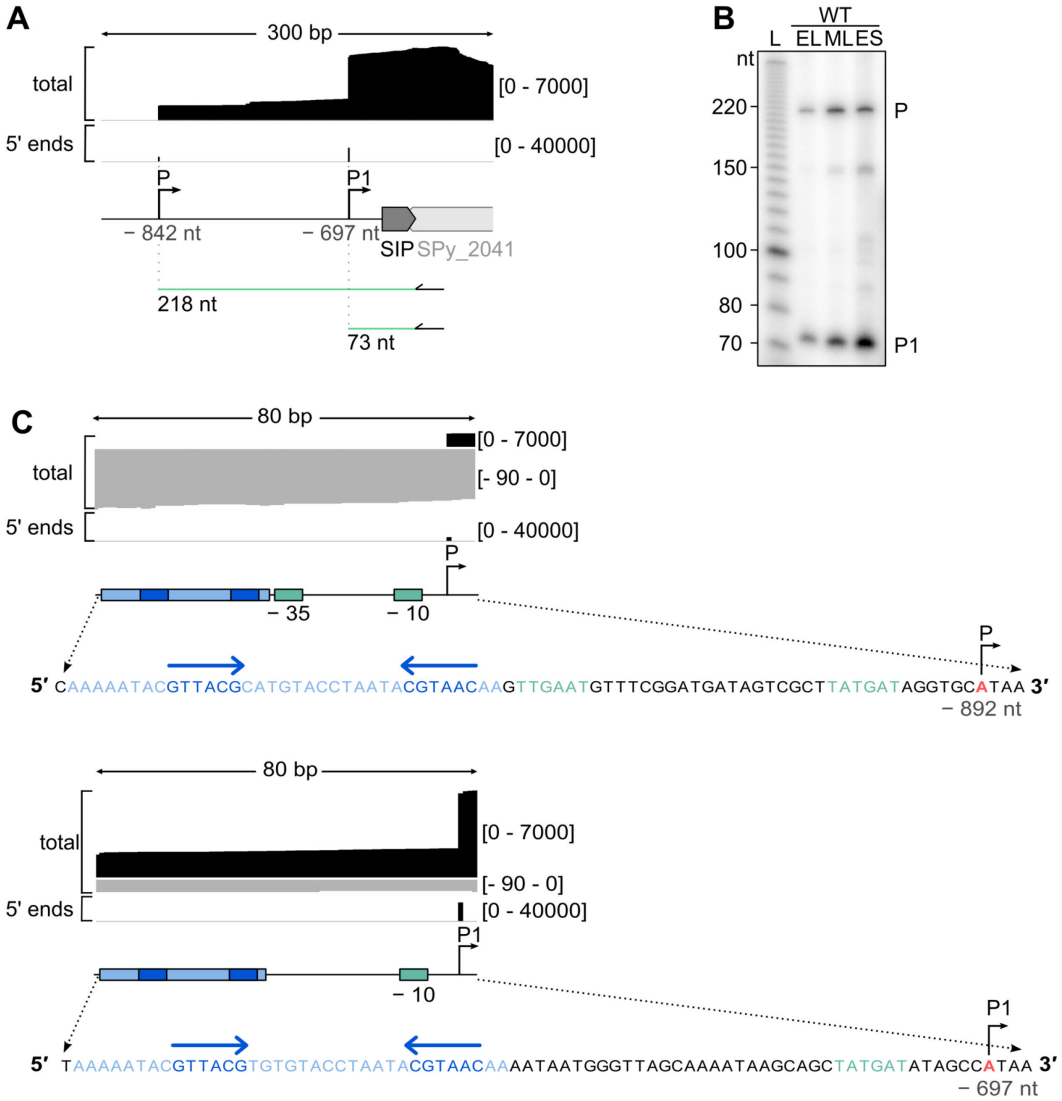


***Figure 12. speB* is transcribed from two promoters.**
A and **C.** Total and 5' end coverage profiles (black from the positive strand and grey for the negative strand) from RNA sequencing. The scales are indicated between brackets. **A.** The portion of *speB* 5' UTR analysed contains the ORF encoding the SpeB inducing peptide (SIP), and a portion of the putative ORF *SPy_2041*. P and P1 (black bent arrows) depict the *speB* transcriptional start sites (TTSs) at – 842 and – 697 nt from the *speB* start codon, respectively. The expected cDNA products obtained in the primer extension analysis and the primer used in the experiment are represented with green lines and a black arrow, respectively. **B.** *speB* TSSs (P and P1) were validated by primer extension performed at early-logarithmic (EL), mid-logarithmic (ML) and early-stationary (ES) phases of growth. The results of one representative primer extension analysis (n=3) are shown. **C.** Zoom on *speB* promoters. The putative – 35 and – 10 motifs (in green) were identified by visual screening. The direct repeats (light blue rectangle) containing inverted repeats (dark blue rectangle and arrows) are part of the putative RopB binding sites located upstream of the P and P1 promoters[301].

3.2.3 RNase Y processes the *speB* mRNA 5' UTR after a guanosine

Besides P and P1, a third *speB* TSS (P2) at – 137 nt relative to the *speB* start codon was previously annotated[301]. However, a study of the P2 activity revealed that this promoter was unable to sustain the expression of a reporter gene[301]. It has been then proposed that this RNA 5' end would derive from the processing of *speB* mRNA 5' UTR by endoRNase(s)[161]. Here, the RNA 5' end at the P2 position was identified by RNA sequencing (Figure 13A), but we could not identify the – 10 and – 35 boxes upstream of P2 by manual screening. To explore the origin of this RNA 5' end, primer extension analyses were performed in the WT and in different RNase deletion mutant strains (Figure 13A and B). The cDNA product corresponding to the RNA 5' end at P2 was absent only in the Δ*rny* strain (Figure 13A and B), demonstrating that RNase Y was responsible for the production of this RNA 5' end. A second RNA 5' end, located 6 nt downstream was also detected by RNA sequencing and generated by RNase Y (Figure 13A-C). In addition, both these RNA 5' ends produced by RNase Y in *speB* mRNA 5' UTR were mapped downstream of a G residue as most of the RNase Y-dependent RNA 5' ends (Figure 11D).

Another RNA 5' end was identified in the *speB* mRNA 5' UTR, 78 nt upstream of the *speB* start codon, by both RNA sequencing and primer extension analyses (Figure 13A-D). It is unlikely that this position corresponded to a TSS because: (i) we did not identify – 10 and – 35 motifs and (ii) the region upstream of this RNA 5' end (up to 623 nt) was previously analysed in a transcriptional reporter assay and it was showed that it did not harbour any promoter activity[301]. Therefore, we hypothesized that this position corresponded to a cleavage site. To pinpoint the RNase responsible for this processing event, primer extension analyses were performed in different RNase deletion mutant strains (Figure 13B). Because this RNA 5' end was present in all the RNase deletion mutant strains tested, the RNase responsible for this processing remains unknown (Figure 13B). It should be notated that RNase J1 is essential for growth in *S. pyogenes*[195], therefore a conditional mutant (constructed by A.-L.L.) was used. In this strain the expression of the *rnjA* gene (coding for RNase J1) is controlled by the tetracycline-inducible promoter (P_{tet})[195,347,348]. However, due to the leakiness of the P_{tet} promoter, *rnjA* is expressed also under repression conditions (*i.e.* in the absence of the inducer) (data not shown). Therefore, we could not conclude whether RNase J1 was involved in this processing event. This RNA 5' end could be produced by RNase(s) for which the deletion mutant was not examined (*e.g.* RNase J2). Alternatively, as RNases often act redundantly, it is possible that more than one RNase can cleave at this position and therefore multiple RNase deletion strains, are required to identify the RNase(s) responsible for the processing.

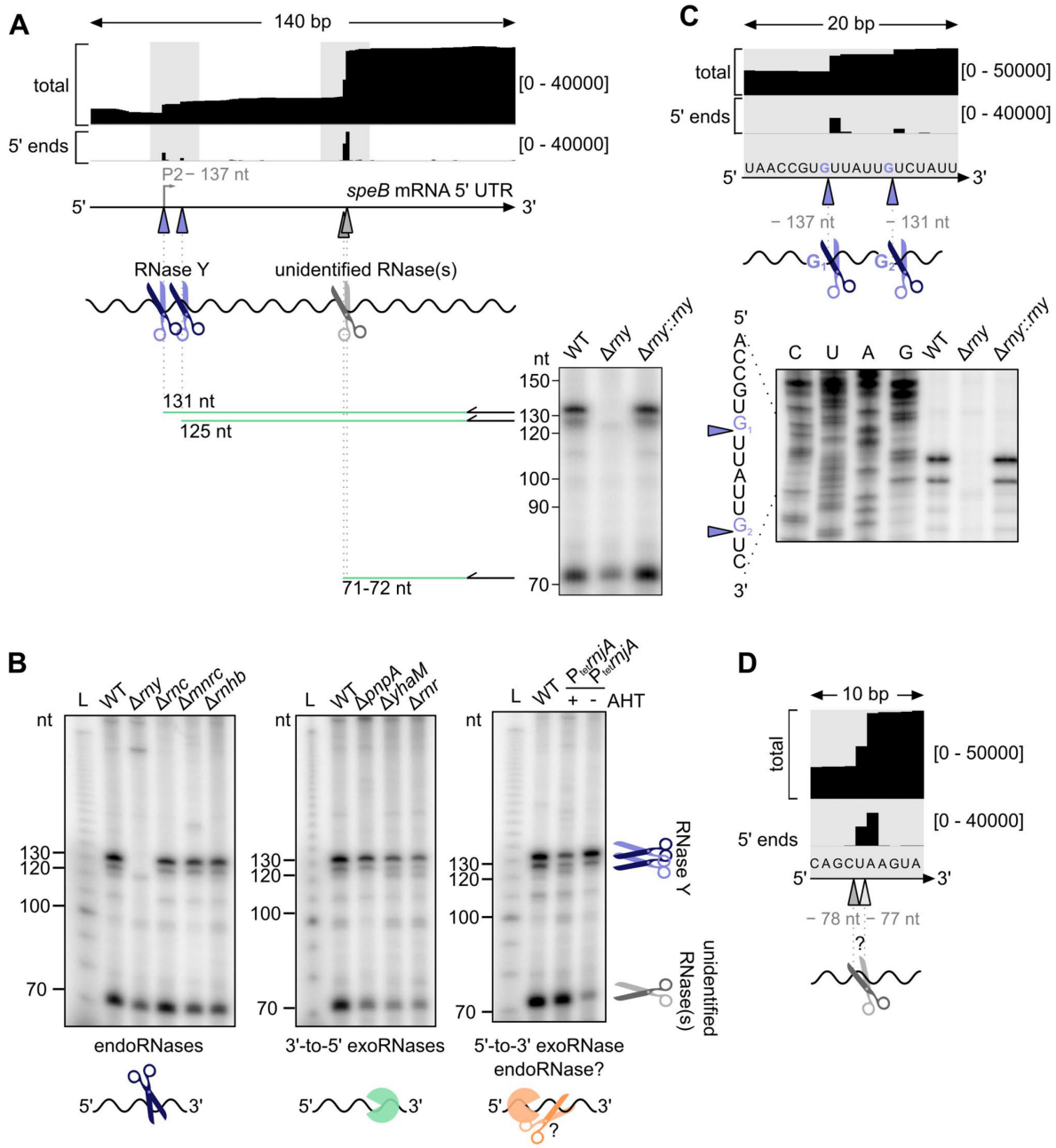


Figure 13. **RNases target the *speB* mRNA 5' UTR.**
A, **C** and **D**. Total and 5' end coverage profiles from RNA sequencing of portions of the *speB* 5' UTR. The coverage scales are indicated between brackets. **A.** The grey rectangles depict the regions that were further analyzed. The RNA 5' ends that correspond to RNase Y (purple scissors) and unidentified RNase(s) (grey scissors) processing sites are indicated with purple and grey arrowheads. **A** and **B**. Mapping of the RNA 5' ends in *speB* 5' UTR by primer extension analyses at early-stationary growth phase. The primers and the length of the expected cDNA products are shown by an arrow and green lines, respectively. The experiments were performed in the WT, *rny* (RNase Y) deletion mutant (Δ*rny*) and chromosomal complemented *rny* deletion mutant (Δ*rny*::*rny*) strains in panel A and also in the *rnc* (RNase III) deletion mutant (Δ*rnc*), *mrnC* (Mini-III) deletion mutant (Δ*mrnC*), *rnhB* (RNase HIII) deletion mutant (Δ*rnhB*), *pnpA* (PNPase) deletion mutant (Δ*pnpA*), *yhaM* (YhaM) deletion mutant (Δ*yhaM*), *rnr* (RNase R) deletion mutant (Δ*rnr*) and *rnjA* (RNase J1) conditional mutant (P_{tet}*rnjA*) strains in panel B. The latter strain was grown in the absence or presence of anhydrotetracycline hydrochloride (AHT), used at a final concentration of 0.1 ng/µL. The results of one representative primer extension analysis (n>3) are shown. S.M. contributed in production of the replicates of the primer extension analysis in the different RNase deletion mutants. **C.** Zoom on the RNase Y processing sites at – 131 and – 137 nt relative to the *speB* start codon validated by

primer extension analyses using a sequencing ladder. The two cleavage sites are located after a guanosine (G). **D**. Zoom on the processing site(s) of unidentified RNase(s) located at – 78/79 nt from *speB* start codon.

In light of the re-annotation of *speB* transcript 5' ends, we concluded that the t3, t4 and t5 transcript isoforms derived from RNase Y processing of the t1 and t2 transcripts. In particular, the t3 transcript isoforms corresponded to the processed RNAs upstream of the RNase Y cleavage in the *speB* mRNA 5' UTR (Figure 14 and Figure 8B). Instead, the t4 and t5 isoforms corresponded to the processed RNA downstream of the RNase Y cleavage site(s) (Figure 14 and Figure 8C). The t5 isoforms that was detected in the Δ*rny* strain (Figure 8C) could be generated by the unidentified RNase(s) processing the *speB* 5' mRNA UTR (Figure 14).

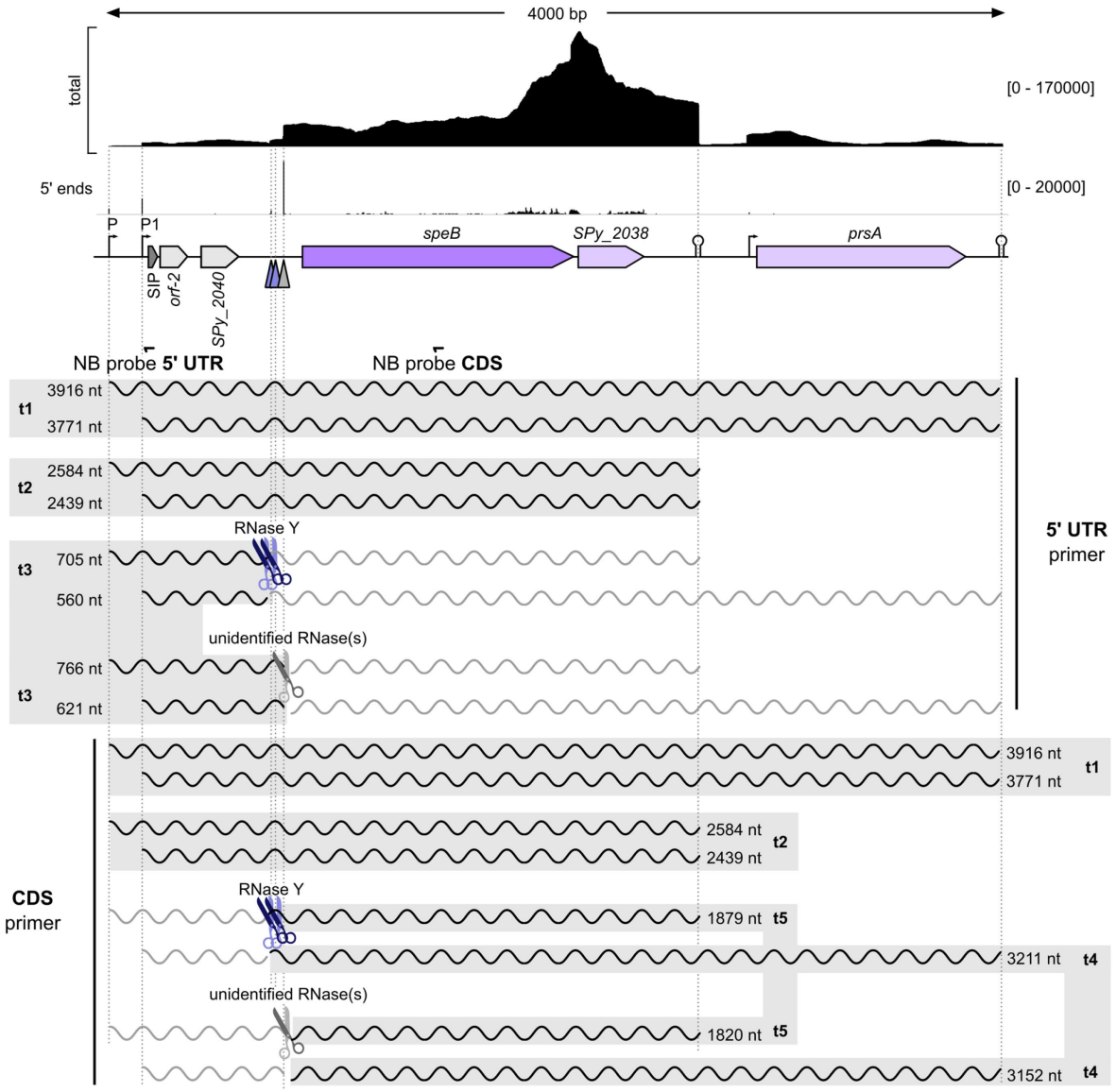


***Figure 14. speB* transcript isoforms detected by Northern blot analyses.**
Total and 5' end coverage profiles from RNA sequencing of *speB* and downstream genes. The coverage scale is indicated between brackets. *speB* 5' UTR comprises small putative ORFs: *SPy*_2040, *orf-2* and the SpeB inducing peptide (SIP). The predicted transcriptional start sites and terminators are depicted in black. *speB*

is co-transcribed together with *SPy_2038* and *prsA* as previously described[331]. The processing positions of RNase Y and unidentified RNase(s) are indicated with purple and grey arrowheads, respectively. The labelled oligonucleotide probes binding to the *speB* 5' untranslated region (5' UTR) (in Figure 8B) and to the *speB* coding DNA sequence (CDS) (in Figure 8C) for the Northern blot analysis (NB probes) are depicted by black arrows below the loci representation. The transcripts that were detectable or undetectable are indicated by black and grey curved lines, respectively. The expected sizes in nucleotides (nt) are reported for each transcript isoform.

3.2.4 Mutation of the G downstream of the RNase Y cleavage sites in *speB* mRNA 5' UTR impedes RNase Y processing

We observed that the majority of the RNA 5' ends produced by RNase Y were mapped after a G (Figure 11F), as also described in *S. aureus*[160], and that RNase Y cleavage sites specifically located in the *speB* mRNA 5' UTR occurred as well downstream of a G (Figure 13C). The role of this nucleotide in RNase Y target recognition remains unknown. To explore whether the G is required for RNase Y processing *in vivo*, *speB* 5' UTR and CDS were cloned under the control of a constitutive promoter on a plasmid, which was transformed in a *speB* (SpeB) deletion mutant (Δ*speB*) strain (Figure 15). The Gs found upstream of the RNase Y processing positions at − 137 nt and – 131 nt from the *speB* start codon were named G_1 and G_2, respectively. These nucleotides were individually or simultaneously substituted into A (Figure 15). Additionally, deletions of 10 or 20 nt upstream and downstream of the RNase Y cleavages were also performed, including the 6 nt between the two cleavage sites (Δ26 and Δ46 nt) (Figure 15). Primer extension analysis was used to assess the impact of these mutations on RNase Y activity in the *speB* mRNA 5' UTR. (Figure 15). The usage of a Δ*speB* strain and of a primer targeting the *speB* CDS in the primer extension analysis, ensures to detect the processing sites occurring on the *speB* mRNA transcribed from the plasmid and not from the bacterial genome. The replacement of G_1 by A_1 strongly impaired RNase Y processing at − 137 nt position (Figure 15; lanes 2 and 5). In contrast, the substitution of G_2 with A_2 completely abolished the processing by RNase Y (Figure 15; lanes 4 and 5). Interestingly, when G_1 was substituted by A_1, an alternative processing site was observed around 2 nt upstream of the cleavage site at – 137 nt (Figure 15; lanes 3 and 5). Sequence analysis of the *speB* mRNA 5' UTR revealed that a G is located 2 nt upstream of the processing at − 137 nt position (Figure 15; see box alternative processing sites). Other alternative processing sites were identified when 10 nt both upstream and downstream of the RNase Y processing positions were deleted (Figure 15; lane 6). These cleavages were mapped approximately 115 nt and 116 nt upstream of the *speB* start codon and they were also detectable when both the G_1 and G_2 were mutated in A (Figure 15; lanes 3, 4 and 5). Of note, a G was also identified at both the − 115 nt and − 116 nt positions, in the *speB* mRNA 5' UTR (Figure 15; see box alternative processing sites). The observation that the alternative cleavage sites likely mapped at a G led to the hypothesis that also these processing

events are performed by RNase Y. This would suggest that, in the absence of the canonical processing site, RNase Y is able to target the *speB* mRNA 5' UTR and to cleave at different Gs located upstream or downstream of the original RNase Y processing sites. However, this hypothesis awaits further evidence, since it is possible that other endoRNase(s) cleave the *speB* mRNA 5' UTR when the processing sites of RNase Y are disrupted. To evaluate whether RNase Y is responsible for the alternative processing sites, these experiments should be performed in the double deletion mutant Δ*speB*Δ*rny* strain. Nonetheless, in this condition, RNase(s) that act redundantly could replace RNase Y activity complicating the analysis of the alternative processing site origin. Finally, no processing events were identified when 20 nt were deleted upstream and downstream of the RNase Y cleavage sites, respectively (Figure 15; lane 7). Therefore, these deletions likely perturbed the *speB* mRNA 5' UTR such that RNase Y and/or other endoRNase(s) were unable to cleave it.

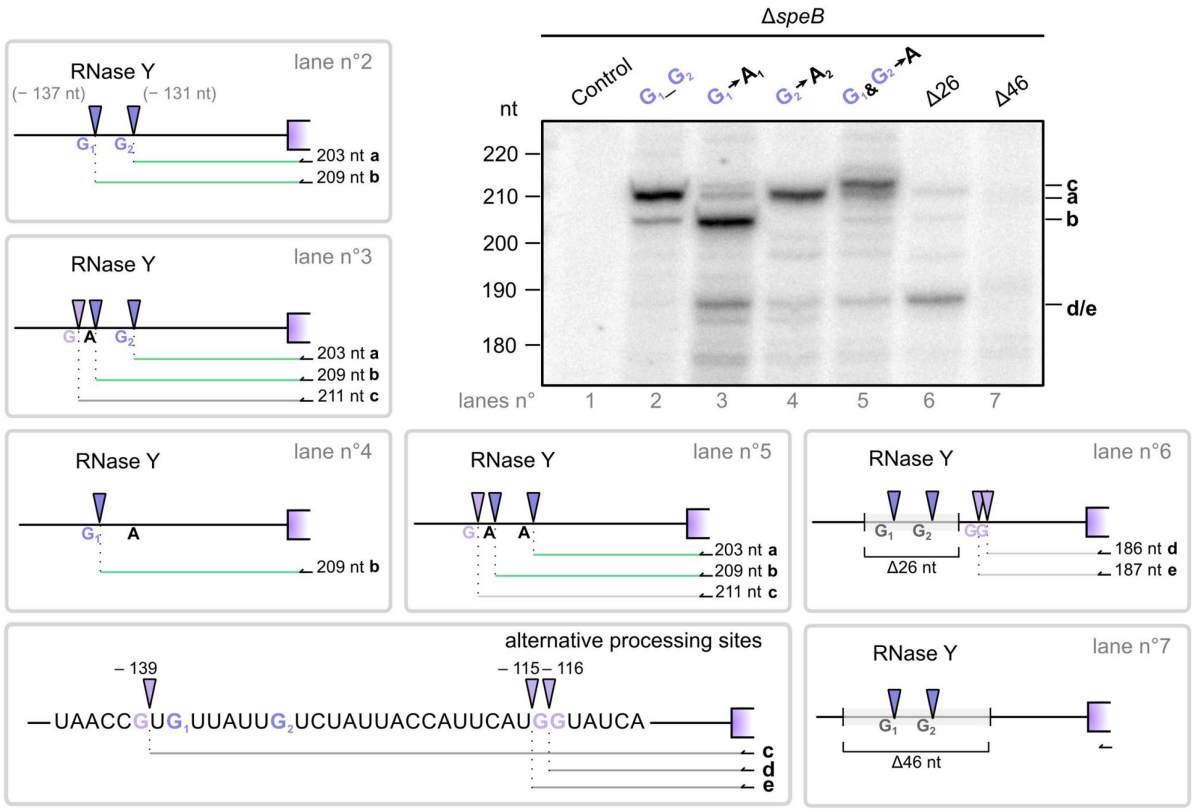


Figure 15. **RNase Y requires a guanosine upstream of the processing site to cleave *speB* mRNA.** Mutational study to assess the importance of the G located upstream of the RNase Y processing sites in the *speB* mRNA 5' UTR. The primer extension analyses were performed at early-stationary growth phase in the *speB* (SpeB) deletion mutant (Δ*speB*) strain, which was transformed with different plasmids, schematically represented in each of the boxes. The plasmids contain the *speB* 5' UTR and CDS under the control of the constitutive promoter P_{gyrA} (see Material & Methods). An empty vector was used as a control in lane n° 1. The Gs upstream of the cleavage sites at – 131 nt and (G_1) – 137 nt (G_2) were mutated into adenosine (A) (lines n° 2, 3, 4 and 5). Alternatively, 10 or 20 nt were deleted upstream and downstream of the RNase Y processing sites, including the 6 nt between the two RNase Y processing positions (lines n° 6 and 7). The primer used in the experiments is shown in black and binds at the beginning of the *speB* CDS. The results

of one representative primer extension analysis (n=3) are shown. The purple arrowheads depict the RNase Y cleavage positions and the light purple arrowheads depict the alternative processing sites detected. The box named "alternative processing sites" contains a representation of the *speB* 5' UTR with the nucleotide sequence from – 145 nt to – 110 nt relative to *speB* start codon and the annotated alternative processing sites.

Overall, these results demonstrated that the G upstream of the RNase Y cleavage sites is crucial for RNase Y processing in *S. pyogenes* and validated the finding of the genome-wide mapping of RNase Y cleavage sites, in which the G upstream of most of the 5' *rny*_ends was predicted to be important for RNase Y activity. On the contrary, RNase Y processing in *S. aureus* was shown to depend on secondary structure elements 6 nt downstream the processing site[169]. Therefore, we next proceeded to investigate the presence of structured RNA regions in proximity of the RNase Y processing events (Figure 16A and B).

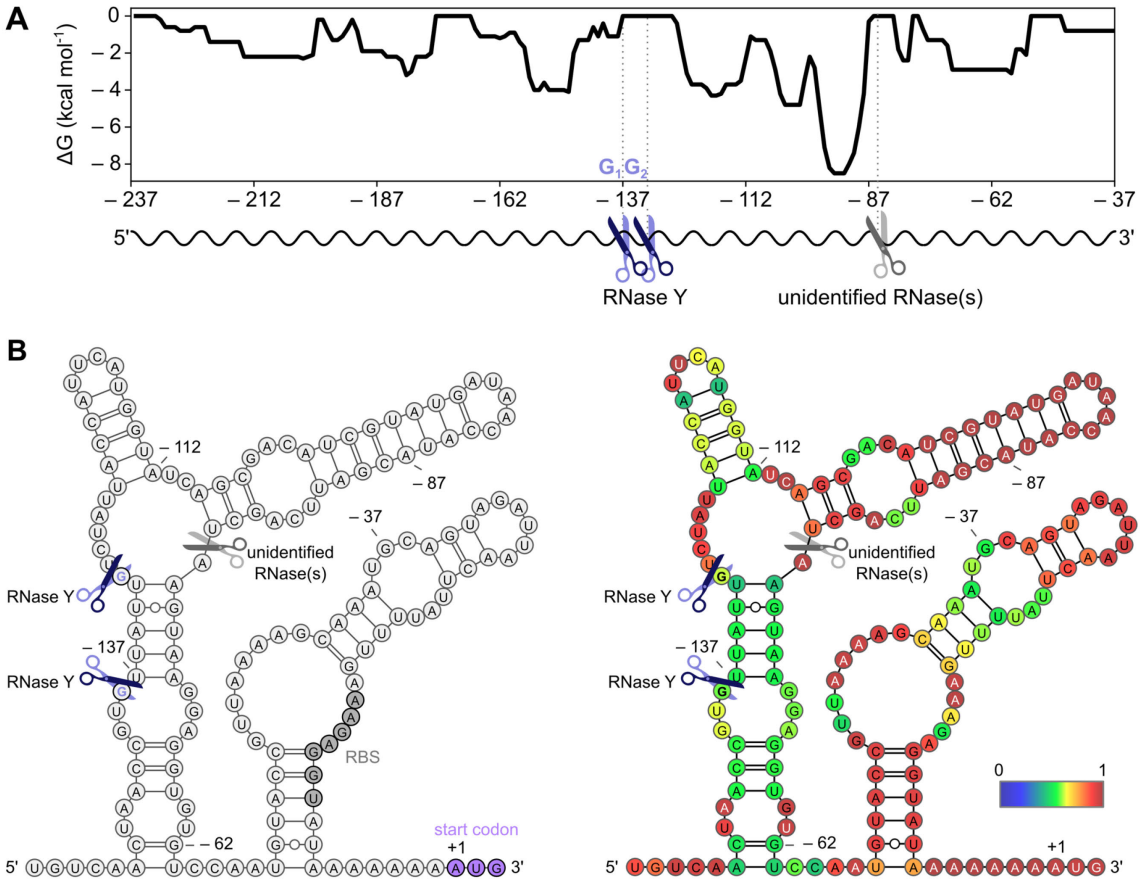


Figure 16. **Prediction of RNA secondary structures in the *speB* mRNA 5' UTR.**
A. Prediction of secondary structures surrounding the RNase Y cleavage sites in the *speB* mRNA 5' UTR. The minimum free energy (ΔG in kcal mol^{-1}) was calculated with a 50 nt sliding window, both 100 nt upstream and downstream of the RNase Y cleavage site at 137 nt before the *speB* start codon. RNase Y (purple scissors) and unidentified RNase(s) (grey scissors) processing sites are indicated. The ΔG calculation was conducted by T.T.R with A.L.R and L.B. **B.** Prediction of the RNA folding of a region of *speB* mRNA 5' UTR performed with RNAfold (see Material & Methods). The free energy of the thermodynamic ensemble is – 31.48 kcal/mol. The *speB* ribosome binding site (RBS) and the *speB* start codon are indicted.

The cleavage positions by RNase Y and unidentified RNase(s) are depicted by purple and grey scissors, respectively. The RNA folding color-scheme is shown on the right. For unpaired regions the colour represents the probability of being unpaired.

At the RNase Y processing sites in *speB* mRNA 5' UTR, an increase of the minimum free energy (ΔG) was observed compared to the surrounding regions (Figure 16A). This is in accordance with the observation that RNase Y cleaves ssRNA regions, as described for the total 5' *rny*_ends mapped in *S. pyogenes* (Figure 11D). A decrease of the ΔG was observed between 40 and 50 nt downstream of the RNase Y processing sites, indicative of a putative RNA structure (Figure 16A, B). Due to the distance of this putative RNA structure from the RNase Y processing in *speB* mRNA 5' UTR, we exclude that the 6 nt ruler and cut mechanism suggested for RNase Y in *S. aureus*[169] and *B. subtilis*[119] is applicable in *S. pyogenes*, at least for the *speB* mRNA processing.

3.2.5 RNase Y produces putative sRNA(s) from *speB* mRNA 5' UTR

A previous RNA sequencing analysis performed in our laboratory uncovered a putative sRNA (Spy_sRNA1699993) expressed from the *speB* mRNA 5' UTR, which I will refer to as *speB*-sRNA in this thesis[359] (Figure 17A). Interestingly, the *speB*-sRNA 5' end corresponded to the RNase Y processing position at –137 nt relative to *speB* start codon. Therefore, it is possible that the *speB*-sRNA biogenesis relied on RNase Y activity in the *speB* mRNA 5' UTR (Figure 17A). According to the previous annotation, the *speB*-sRNA 3' end would coincide with the processing of the unidentified RNase(s), resulting in an approximately 59 nt long-sRNA[359]. However, the maximal length of the reads in this previous study was 85 nt, therefore the RNA 3' ends of longer RNAs could not be determined accurately. Here, the new RNA sequencing data revealed that the *speB*-sRNA 3' end is likely located downstream of the processing event by the unidentified RNase(s), at – 27/– 28 nt from the *speB* start codon, thus the *speB*-sRNA is approximately 110 nt long (Figure 17A). Consequently, the processing by the unidentified RNase(s) is located between the *speB*-sRNA 5' and 3' ends. The *speB*-sRNA 5' and 3' ends were validated by circularization of the RNA after self-ligation followed by reverse transcriptase-polymerase chain reaction (Figure 17B and see Materials & Methods). Most of the self-ligated products harboured the RNA 5' end at – 137 nt, coinciding with the first RNase Y processing site in the *speB* mRNA 5' UTR, and the RNA 3' ends were located between – 25 and – 28 nt from the *speB* start codon (Figure 17B; Table A3). The observation that the *speB*-sRNA 3' end did not map at one exact nucleotide, was consistent with the RNA sequencing data, in which several RNA 3' ends were identified at consecutive positions (Figure 17A). This could be indicative of the activity of a 3'-to-5' exoRNase

that trims the target RNA and do not stop precisely at one position, but instead on successive nucleotides[373–375].

The production of the *speB*-sRNA was analysed over growth using Northern blot analyses in both the WT and Δ*rny* strains (Figure 17C and D). The *speB*-sRNA generation was RNase Y-dependent and occurred only at early-stationary growth phase (Figure 17C and D). The analysis of the transcript stability unveiled that the *speB*-sRNA accumulated over time likely because, although transcription was repressed by rifampicin, the *speB*-sRNA continued to be generated by the RNase Y processing of the *speB* mRNA 5' UTR (Figure 17C). To accurately evaluate the *speB*-sRNA stability, both transcription and RNase Y activity should be inhibited. Two *speB*-sRNA isoforms were detected using a labelled oligonucleotide probe hybridizing to the region of *speB* mRNA 5' UTR between RNase Y and the unidentified RNase(s) processing sites. These two isoforms could correspond to the products resulting from RNase Y processing and ending at the predicted RNA 3' end (Figure 17A and C). When the *speB*-sRNA region between the cleavage site by the unidentified RNase(s) and the predicted *speB*-sRNA 3' end was probed in the Northern blot analysis, a 50 nt-long RNA was detected and it likely resulted from the unidentified RNase(s) processing (Figure 17A and D). The transcript isoforms upstream of the unidentified RNase(s) processing event were not observed, indicating that were likely degraded (Figure 17A and D).

While we determined that the *speB*-sRNA isoform 5' ends were generated by RNase Y processing, the origin of the *speB*-sRNA 3' end is unclear. We did not find any predicted intrinsic terminator upstream of the *speB* CDS, which would explain the generation of the *speB*-sRNA 3' end. It is possible that after transcription termination further down in the *speB* CDS or an endoRNase cleavage, the RNA 3' end is processed by 3'-to-5' exoRNase(s), which would give rise to the mature *speB*-sRNA 3' end. Several attempts to overexpress the *speB*-sRNA from a vector, in which a terminator was inserted downstream of the *speB*-sRNA sequence, were unsuccessful as the production of the *speB*-sRNA 3' end was impaired (data not shown). This observation indicates that the terminator structure could interfere with the mechanism of the *speB*-sRNA 3' end generation. For instance, if the *speB*-sRNA 3' end would be generated by a 3'-to-5' exoRNase, as hypothesized above, the terminator structure could block the 3'-to-5' exoRNase trimming.

These results indicate that a possible functional consequence of the RNase Y processing of the *speB* mRNA 5' UTR is the production of a putative sRNA, which could exert potentially a regulatory role in gene expression.

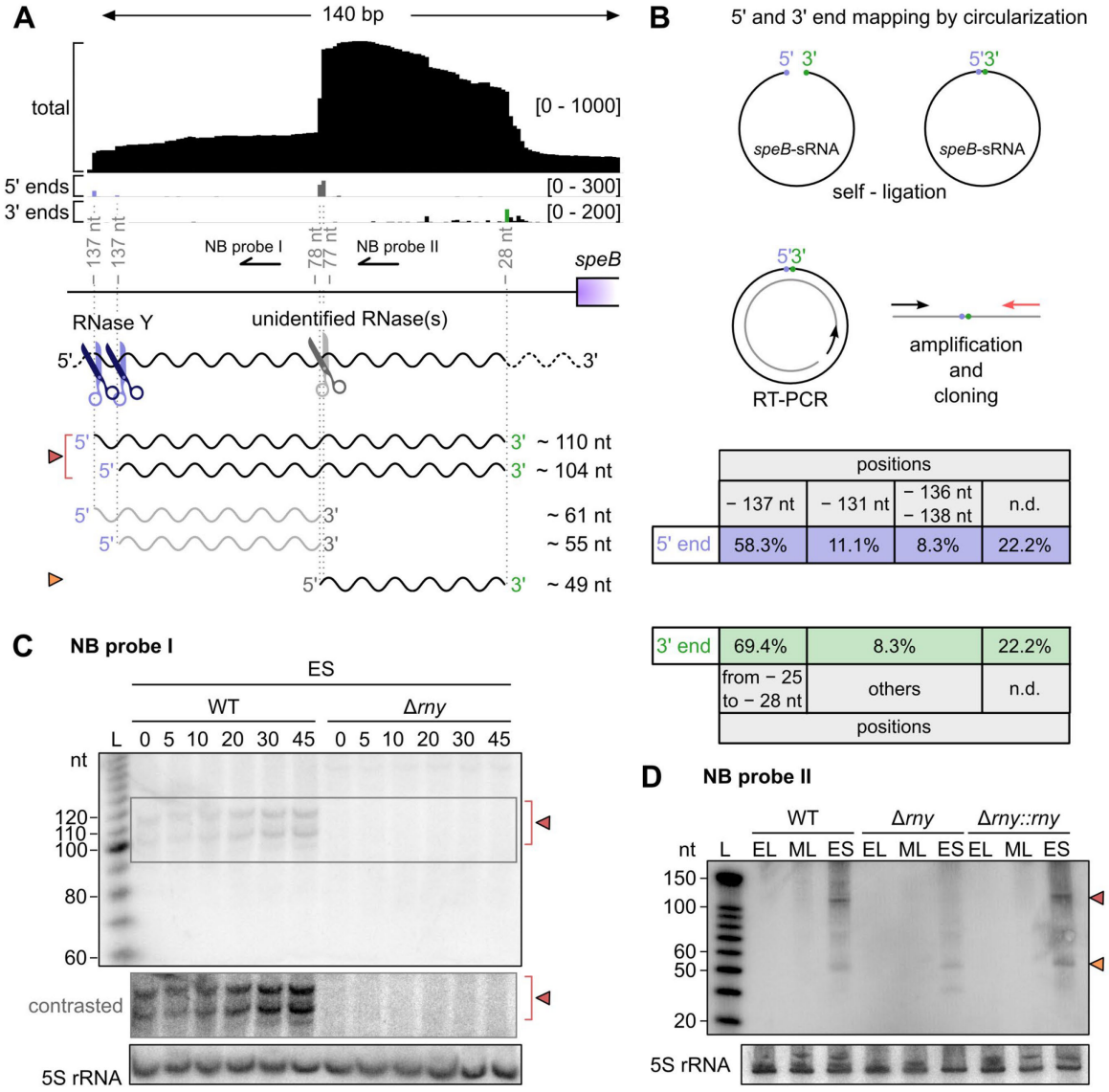


Figure 17. **Putative sRNA(s) derive from the RNase Y processing of the *speB* mRNA 5' UTR.** **A.** Total, 5' end and 3' end coverage profiles from RNA sequencing of the putative sRNA (Spy_sRNA1699993) previously identified in the *speB* mRNA 5' UTR[359], named here *speB*-sRNA. The coverage scale is indicated between brackets. The purple and grey bars denote the processing positions by RNase Y (purple scissors) and unidentified RNase(s) (grey scissors), respectively. The green bar denotes the putative sRNA 3' end. The numbers indicate the location of the positions described above relative to the *speB*-sRNA start codon. The labelled oligonucleotide probes used in the Northern blot analyses (in panels C and D) and the expected detectable transcripts are shown below the locus representation. **B.** Simultaneous mapping of the *speB*-sRNA 5' and 3' ends by circularization of RNA after self-ligation followed by Reverse Transcriptase-Polymerase Chain Reaction (Circ-RT-PCR). On the top, schematic representation of the procedure (see also Material & Methods). On the bottom, the identified *speB*-sRNA 5' and 3' ends are reported together with the proportion of clones in which these RNA ends were mapped. The sequenced *speB*-sRNA self-ligated region is reported in Table A3. The positions are referred to the *speB* start codon (see panel A). The n.d. (not determined) indicates the cases in which the mapping of the RNA 5' and 3' ends was unsuccessful. **C.** and **D.** Study of *speB*-sRNA production by Northern blot analyses, the labelled oligonucleotide probes used and the expected detectable RNAs are reported in panel A. The arrows at the side of the blot indicates the bands that could correspond to the predicted *speB*-sRNA isoforms shown in panel A. The 5S rRNA was used as a loading control. The results of one representative Northern blot analysis (n=3) are shown. **C.** The *speB*-RNA stability was analysed at early-stationary (ES) phase of growth in the WT and the *rny* (RNase Y) deletion mutant (Δ*rny*) strains. The minutes after stopping the transcription

upon addition of the rifampicin are indicated. A contrasted portion of the blot is shown below the full blot. **D.** The pattern of *speB*-sRNA production was analysed at early-logarithmic (EL), mid-logarithmic (ML) and early-stationary (ES) growth phases in the WT, Δ*rny* and in the chromosomal complemented *rny* deletion mutant (Δ*rny::rny*) strains. K.H. contributed in the replicate generation of the *speB*-sRNA Northern blot analyses.

3.2.6 Conclusions II

The genome-wide mapping of the RNA 5' ends produced by RNase Y revealed that RNase Y recognizes a G located upstream of the processing sites. We re-annotated the *speB* mRNA 5' UTR and observed that it was highly processed by RNases, including RNase Y that cleaved at two different locations after a G. The *speB* mRNA 5' UTR was used as a model target to dissect the importance of the G for RNase Y activity. This work demonstrates for the first time that this nucleotide is essential for RNase Y activity, because the substitution of the G with A either inhibited or strongly impaired the processing of the *speB* mRNA 5' UTR. It is unclear what is the functional consequence of the RNase Y processing on the SpeB production. We propose that RNase Y activity on the *speB* mRNA 5' UTR leads to the generation of short RNA fragments. From the observation that the generation of these fragments is growth phase-dependent and requires RNase Y, we predict that these RNA fragments might correspond to regulatory sRNA(s).

3.3 Determining the fate of RNase Y-targeted RNAs

3.3.1 Mapping of RNA 3' ends in the WT and Δ*rny* strains

The cleavage by an endoRNase results in the production of two RNA products, harbouring at the termini the newly generated 5' and 3' ends, respectively. To provide a more comprehensive picture of the RNase Y targetome, the RNase Y-dependent RNA 3' ends were also mapped (Figure 18A, B and Table A4), as previously described for the RNA 5' ends (see Figure 11 and Table A2). The RNA 3' ends that were more abundant in the WT and the Δ*rny::rny* strains than in the Δ*rny* strain were annotated as 3' *rny*_ends (Figure 18B). The abundance of 130 RNA 3' ends depended on the presence of RNase Y (Table A4). As for the 5' *rny*_ends, the structural and sequence context around the identified 3' *rny*_ends were examined and strikingly, the latter harboured distinct characteristics compared to the former. First, a decrease in the ΔG, indicative of a structured RNA region, was observed upstream of the 3' *rny*_ends (Figure 18C), but not for the 5' *rny*_ends (Figure 11D). Second, we did not identify a G or any other conserved sequence in proximity of the 3' *rny*_ends, by examining the sequence alignment of 10 nt upstream and downstream of the 3' *rny*_ends (Figure 18C and see Material & Methods).

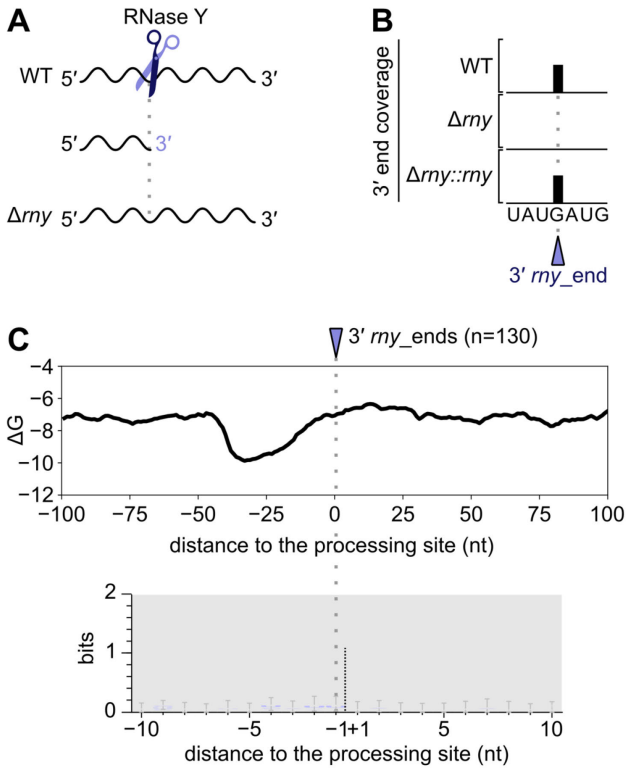


Figure 18. **Mapping of the RNase Y-dependent RNA 3' ends.**
A. Example of an RNA molecule which is processed by RNase Y (scissors) producing a new RNA 3' end detectable in the WT strain but not in the *rny* (RNase Y) deletion mutant (Δ*rny*) strain. **B.** Representation of RNA sequencing 3' end profile. The RNA 3' end more abundant in the WT strain and in the complemented

rny deletion mutant (Δ*rny::rny*) strain than in the Δ*rny* strain were annotated as 3' *rny*_end (purple arrowhead). The "UAUGAUG" represents a random sequence cleaved by RNase Y. **C.** Structure (on the top) and sequence (on the bottom) conservation of the 3' *rny*_ends. The minimum free energy (ΔG, in kcal/mol^{-1}) was calculated at each nucleotide using a 50 nt sliding window, 100 nt upstream and downstream of the 3' *rny*_ends. The logo was generated by aligning 10 nt each side of the 130 5' *rny*_ends. The sequence and structure analyses at the 3' *rny*_ends were conducted by T.T.R. with A.L.R, A.-L.L. and L.B.

3.3.2 RNA 3' ends produced by RNase Y are trimmed by 3-to-5' exoRNases

The analysis of the 3' *rny*_end and 5' *rny*_end positions revealed that these RNA ends were never mapped at consecutive nucleotides (Tables A2 and A4), suggesting that they did not derive from the same processing event. Based on this observation, we concluded that RNase Y processing always led to RNA degradation, likely performed by exoRNases, of one or both the generated RNA products (Figure 19A).

The identified 3' or 5' *rny*_ends were mapped either at consecutive nucleotides or at one specific nucleotide (Figure 19B, Tables A2 and A4). When the *rny*_ends were found at successive nucleotides, only one position corresponding to the one with the highest proportion of ends was annotated (see Material & Methods) and defined as "stepped" (S)-RNA end (Figure 19B). In contrast, the RNA ends identified at only one nucleotide were named "unique" (U)-RNA end (Figure 19B). We observed that the 3' *rny*_ends were predominantly S-RNA ends, while the 5' *rny*_end were in most of the cases U-RNA ends (Figure 19C, Tables A2 and A4).

Based on the observations that (i) the 3' *rny*_ends, but not the 5' *rny*_ends, were located downstream of a secondary structure, (ii) the conservation for the G was not observed at the 3' *rny*_ends, and that (iii) the 3' *rny*_ends were mainly classified as S-RNA ends, we hypothesized that the 3' *rny*_ends actually derived from the trimming of 3'-to-5' exoRNase upon RNase Y processing (represented in Figure 19D). As mentioned before, 3'-to-5' exoRNases, during RNA trimming usually do not stop at one specific position but rather in a region spanning few consecutive nucleotides[373–375], therefore explaining the 3' *rny*_ends "stepped" profile observed in the RNA sequencing.

The trimming positions of three 3'-to-5' exoRNases: YhaM, PNPase and RNase R were previously identified and characterized by our laboratory in *S. pyogenes*[204] (see Figure 29) . The RNA 3' ends that were more abundant in one of the three 3'-to-5' exoRNase deletion mutant strains (*i.e.* Δ*yhaM*, Δ*pnpA* or Δ*rnr*) than in the WT strain corresponded to the positions where the 3'-to-5' exoRNases started the degradation (Figure 19E). Conversely, the RNA 3' ends that were more abundant in the WT strain than in the Δ*yhaM*, Δ*pnpA* or Δ*rnr* strains were designated as the position where the 3'-to-5' exoRNase stopped the trimming[204] (Figure 19E). These available data were used

to study the RNA degradation initiated by RNase Y processing and performed by the three 3'-to-5' exoRNases.

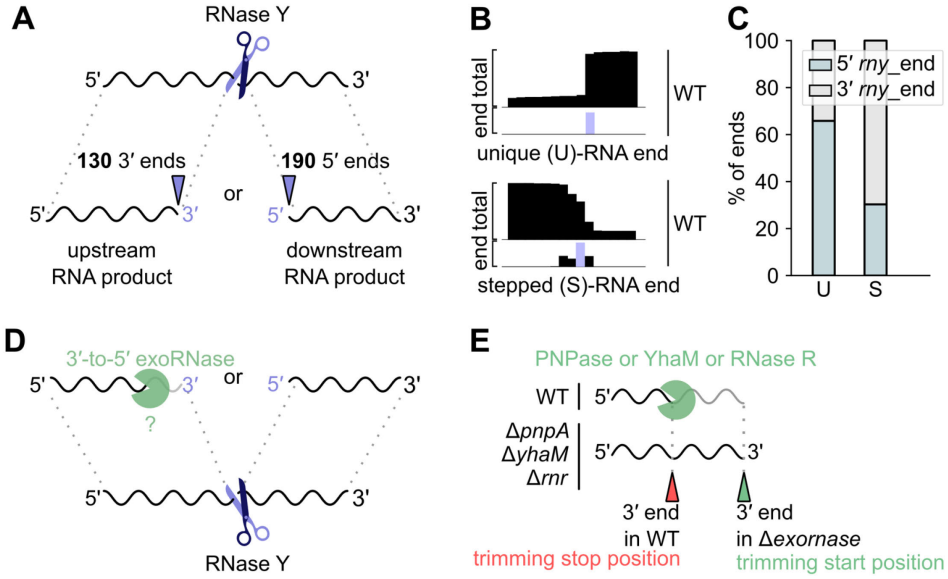


Figure 19. **RNase Y processing initiates RNA degradation.**
A. RNase Y (scissors) processing generates two RNA products, one upstream and one downstream relative to the RNase Y cleavage site. The two RNA products derived from one processing events were never detected together. **B.** Representation of total and 5' or 3' end coverage profiles from RNA sequencing. The *rny*_ends mapping at one single position or at consecutive positions were defined as "unique" (U)-RNA ends or "stepped" (S)-RNA ends, respectively. **C.** Bar plot depicting the portion of 5' *rny*_ends and 3' *rny*_ends that were classified as U- and S-RNA ends. The analysis of the U- and S-RNA ends was performed by A.L.R., A.-L.L. and L.B. **D.** Upon RNase Y (scissors) processing, the RNA product upstream of the processing site is likely further trimmed by 3'-to-5' exoRNase(s) ("pacman" symbol). **E.** The 3'-to-5' exoRNases (PNPase, YhaM and RNase R) trim the RNA 3' end from the start position (corresponding to the 3' end present in the Δ*exornase*, green arrowhead) to the stop position (corresponding to the 3' end present in the WT strain, red arrowhead). The RNA 3' ends produced by 3'-to-5' exoRNases in *S. pyogenes* were previously identified by our laboratory[204].

3.3.3 PNPase and YhaM trim the RNA products upon RNase Y processing

To explore whether the identified 3' *rny*_ends derived from trimming of the 3'-to-5' exoRNases, as hypothesized above, the RNase Y targetome was compared with the PNPase, YhaM and RNase R targetomes[204]. In other words, the genomic locations of the 3' *rny*_ends were compared with the ones of the 3'-to-5' exoRNase trimming start and stop positions, previously identified (Figure 20, see Materials & Methods). This comparative analysis revealed that 58% of the annotated 3' *rny*_ends corresponded either to 3'-to-5' exoRNase trimming start or stop positions, with most of the trimming positions resulting from PNPase activity (Figure 20A). Although the remaining 42% was not associated to any trimming position (Figure 20A), this portion of 3' *rny*_ends was also likely subjected to 3'-to-5' exoRNase activity, as no preference for

a G was observed at the 3' *rny*_ends. It is possible that either these 3' *rny*_ends were trimmed by several 3'-to-5' exoRNases acting redundantly or that they were targeted by other unidentified 3'-to-5' RNases

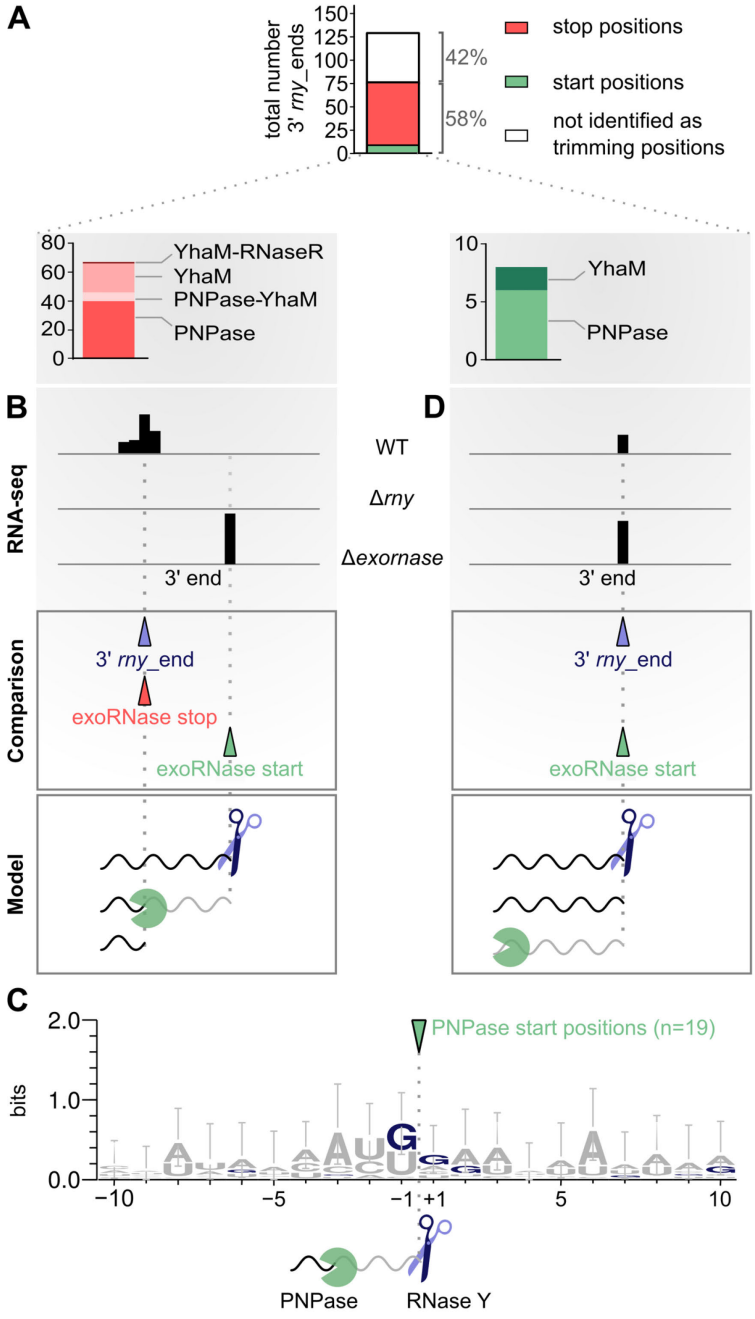


Figure 20. **PNPase and YhaM trim the RNAs processed by RNase Y.**
A. The bar plot on the top shows the portion of RNA 3' ends (3' *rny*_ends) that: (i) were not associated to 3'-to-5' exoRNase processing positions (top in white), (ii) corresponded to the 3'-to-5' exoRNase trimming stop positions (middle in red) and (iii) corresponded to the 3'-to-5' exoRNase trimming start positions (bottom in green). The bar plots at the bottom show the portions of trimming stop (on the right) and trimming start (on the left) corresponding to 3' *rny*_ends, which were generated by one 3'-to-5' exoRNase (PNPase, YhaM or RNase R) or dependent on two 3'-to-5' exoRNases. **B.** and **D.** The 3' *rny*_ends, from the RNase Y targetome, were compared with the 3'-to-5' exoRNase trimming stop (right) and start (left) positions. RNA-seq: schematic representation of the 3' end coverage profile from RNA sequencing; Comparison: RNA ends were included in the comparative analysis. The RNA ends more abundant in the

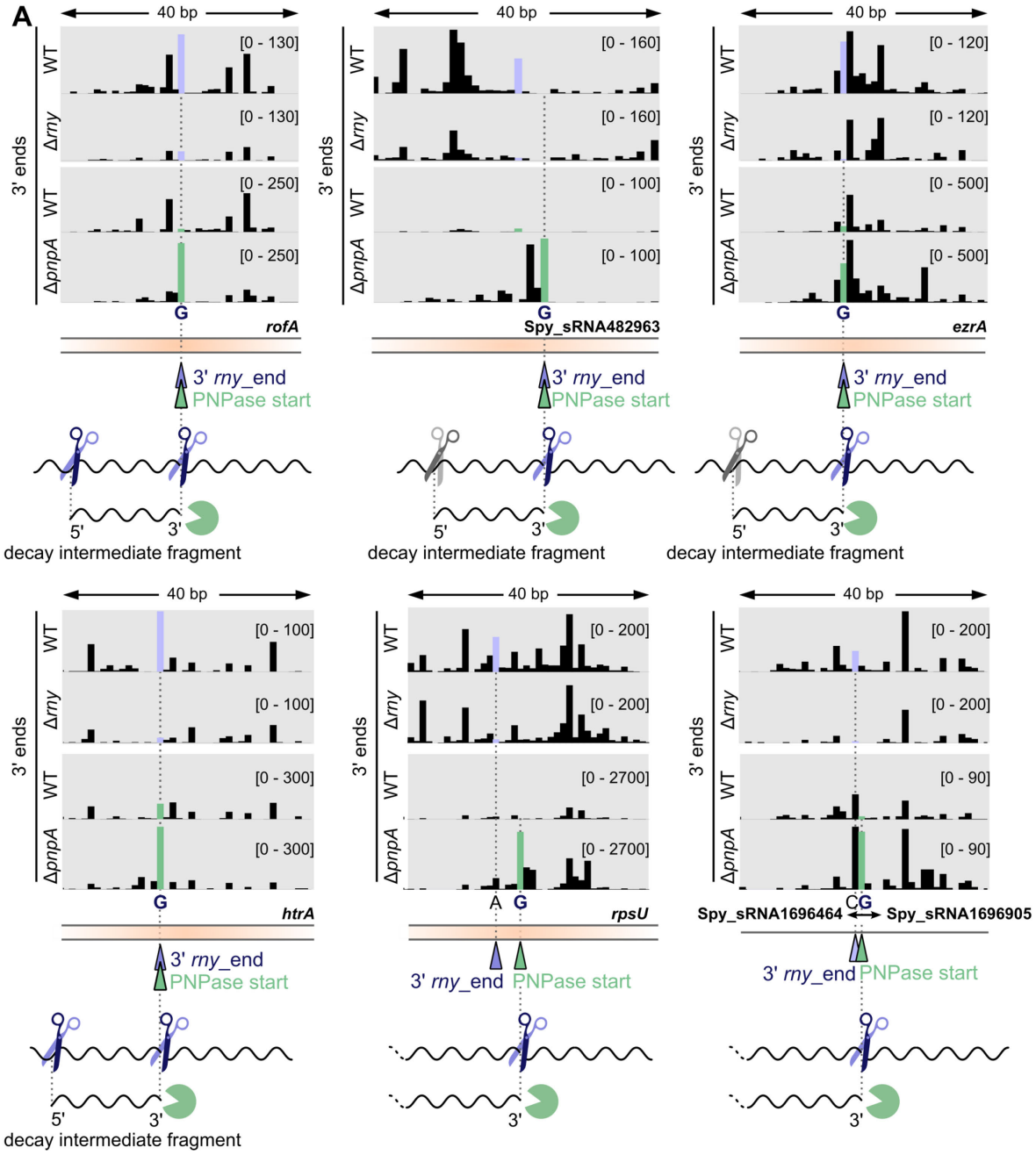
WT strain than in the *rny* (RNase Y) deletion mutant (Δ*rny*) strain are depicted with purple arrowheads (3' *rny*_ends) and the 3'-to-5' exoRNase trimming start and stop positions are indicated with green and red arrowheads, respectively (see Figure 19). The targetome comparisons were performed by T.T.R. and the analyses were conducted by L.B, A.L.R, and A.-L.L. **B.** The 3' *rny*_ends corresponded to the 3'-to-5' exoRNase trimming stop positions (Table A5). Model: the 3'-to-5' exoRNase starts trimming the target RNA at the RNase Y initial processing position and stops before the RNA termini. **D.** The 3' *rny*_ends corresponded to 3'-to-5' exoRNase trimming start positions (Table A6). Model: a 3'-to-5' exoRNase digests the RNAs cleaved by RNase Y. A subset of the 3'-to-5' exoRNase start positions is detectable in the WT strain, suggesting the 3'-to-5' exoRNase does not degrade the whole pool of RNAs. **C.** Alignment of the sequences 10 nt upstream and downstream of the 19 trimming start positions identified downstream of the 3' *rny*_ends corresponding to the PNPase trimming stop positions. The sequence alignment around the PNPase trimming start positions was executed by L.B. and A.L.R.

The comparative analysis of the 3' *rny*_ends with the trimming stop positions revealed the RNAs that were processed by RNase Y and subsequently trimmed by a 3'-to-5' exoRNase, which stopped the degradation before the RNA 3' terminus (Figure 20A, B and Table A5). In few cases, the RNAs were trimmed by more than one 3'-to-5' exoRNase, either by YhaM and RNase R or by YhaM and PNPase (Figure 20A and Table A5).

Since these 3' *rny*_ends corresponded to the position where the 3'-to-5' exoRNase stopped digesting the RNAs we searched for the trimming positions that should coincide with the initial RNase Y processing event (Figure 20B). Therefore, to provide a more accurate annotation of the RNase Y processing sites, we inspected the trimming start positions located downstream of the 3' *rny*_ends corresponding to the trimming stop positions (Figure 20B and Table A6). We identified 19 and 5 trimming start positions of PNPase and YhaM, respectively (Table A6). From the alignment of the sequences surrounding the 19 trimming start positions, we observed a preference for a G in proximity of the PNPase start positions (Figure 20C and Table A6). From this observation we deduced that the location where PNPase started the trimming could correspond to the initial RNase Y processing site. Nine of these trimming start positions were not found exactly at a G, but 1 or 2 nt before. It is possible that this shift from the G was a consequence of YhaM trimming, which removes on average 3 nt from the 3' ends of the majority of all the RNAs in *S. pyogenes*. For YhaM, 4 trimming start positions were identified at a G (Table A6) and one was mapped at an A, which coincided to a PNPase stop position in *SPy_0316* (Tables A5 and A6) which will be further analysed in section 3.3.4.

The comparison analysis described above allowed to identify the RNAs cleaved by RNase Y and already trimmed by 3'-to-5' exoRNase. Next, to detect the RNAs which were not yet targeted by a 3'-to-5' exoRNase, the 3' *rny*_ends were compared to the 3'-to-5' exoRNase trimming start positions and we identified six and two 3' *rny*_ends corresponding to PNPase and YhaM trimming start positions, respectively (Figure 20A and D, Table A7). Although, these RNA 3' ends were targeted by the 3'-to-5' exoRNases, they were still detectable in the WT strain, suggesting that

a subset of each RNA was not degraded yet by the 3'-to-5' exoRNases. PNPase trimming start positions were either perfectly matching the 3' *rny*_ends, located after a G, or they were few nucleotides apart (Figure 21A). In the latter case, PNPase start positions were mapped after a G, thus likely pinpointing the initial RNase Y cleavage position. For some of the RNAs (*rofA*, SPy_sRNA482963, *ezrA* and *htrA*), the 3' *rny*_ends and PNPase trimming start positions were at the RNA termini of previously observed decay intermediate fragments, which were completely degraded by PNPase[204] (Figure 21A). Instead, the 3' *rny*_ends and PNPase trimming start in *rpsU* and in the intergenic region between Spy_sRNA1696464 and Spy_sRNA1696905 did not correspond to the 3' termini of decay intermediate fragments. In these cases, the PNPase trimming



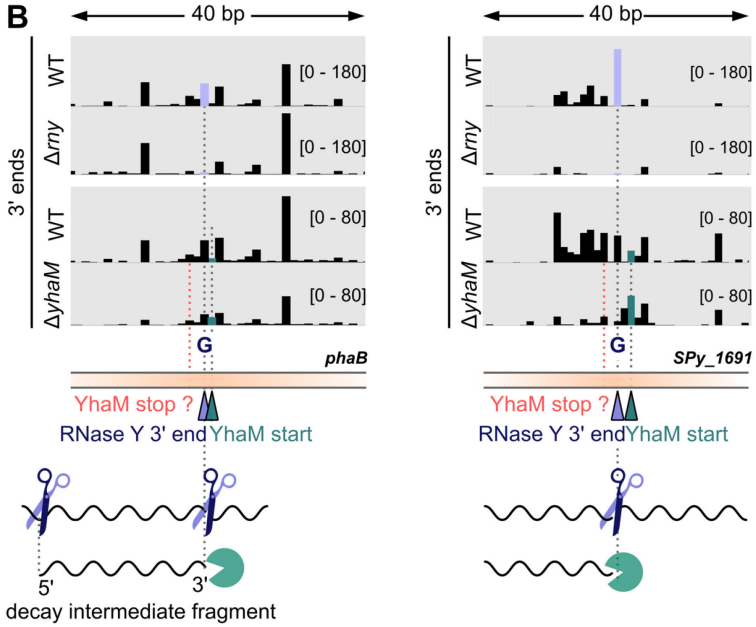


Figure 21. **RNase Y-dependent RNA 3' ends correspond to PNPase or YhaM trimming starts.** **A.** and **B**. RNA 3' end profiles from RNA sequencing in WT, *rny* (RNase Y) deletion mutant (Δ*rny*) and *pnpA* (PNPase) deletion mutant (Δ*pnpA*) strains and in panel B also in the *yhaM* (YhaM) deletion mutant (Δ*yhaM*) strain, showing the 3' *rny*_ends corresponding to PNPase (in panel A) or YhaM (in panel B) trimming start positions (see Figure19D). The coverage scales are indicated between brackets. As both the 3' *rny*_end and the 3'-to-5' exoRNase trimming start positions were present in the WT strain, we deduced that not all the RNA 3' ends were targeted by the 3'-to-5' exoRNases (green "pacman" symbols). The 3' *rny*_ends were often the RNA termini of decay intermediate fragments. The fragment 5' ends were produced by RNase Y (purple scissors) or by an unidentified endoRNase (grey scissors). For the RNAs targeted by YhaM, the trimming stop positions were not identified. However, downstream of the 3' *rny*_ends matching the YhaM trimming start positions, we observed, by manual screening, RNA 3' ends which were more abundant in the WT strain than in the Δ*yhaM* strain, suggesting that these 3' ends could correspond to putative YhaM stop positions not retrieved by the bioinformatic analysis used to identify the RNase processing sites.

stop positions were not identified and the RNAs were likely fully degraded by this exoRNase (Figure 21A). The two YhaM trimming start positions were located at an A, 1 or 3 nt upstream of the 3' *rny*_ends. The locations where YhaM stopped the trimming of these two RNAs were not identified (Figure 21B). However, by manual screening we observed upstream of the YhaM trimming start positions RNA 3' ends which were more abundant in the WT strain than in the Δ*yhaM* strain, indicating that these RNA 3' ends could correspond to putative YhaM trimming stop positions that were not identified in the previous analysis[204] (Figure 21B).

3.3.4 RNA structures protect RNase Y targets from PNPase degradation

Based on the comparative analysis of the 3' *rny*_ends with the trimming stop positions, we could observe that upon RNase Y processing the RNA target was incompletely degraded by a 3'-to-5' exoRNase. As mentioned already, YhaM trims on average 3 nt after terminator regions or

after endoRNase processing[204], explaining the incomplete degradation observed for the RNAs targeted by RNase Y and subsequently trimmed by YhaM (Figure 20A and B). To elucidate the reason why PNPase stopped trimming the RNAs produced by RNase Y, we inspected for secondary RNA structures (see Material & methods) that could resist the enzyme exoribonucleolytic activity. We first compared the ΔG upstream of the RNase Y-depending RNA 3' ends that corresponded or not to PNPase stop positions (Figure 22A). As previously observed for the total 3' *rny*_ends (Figure 18C), the sequences upstream of both the 3' *rny*_ends trimmed and not trimmed by PNPase were structured. However, the decrease in ΔG upstream of the 3' *rny*_ends trimmed by PNPase was stronger than the one observed before the 3' *rny*_ends not targeted by this exoRNase, suggesting that the structures before these 3' RNA ends, trimmed by PNPase, were more stable (Figure 22A).

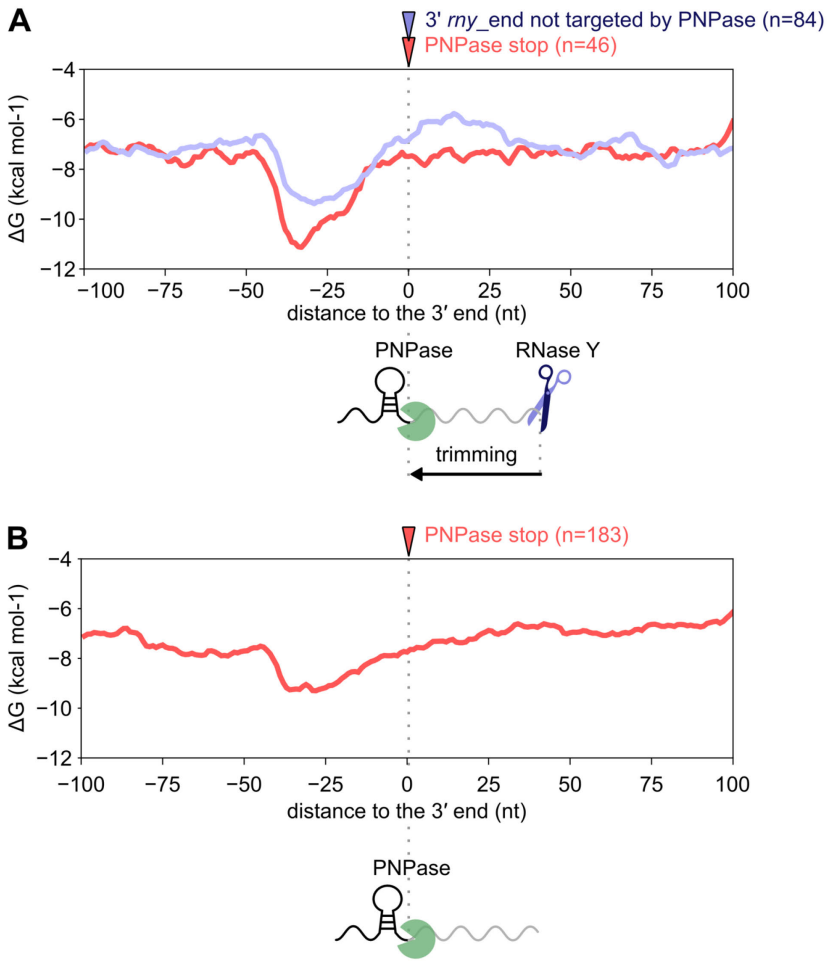


Figure 22. **PNPase degrades the RNAs after RNase Y processing until secondary structures.**
A. and **B.** Structure conservation analysis by calculating the minimum free energy (ΔG in kcal mol^{-1}) in a window of 200 nt around the positions of interest. **A.** Structure conservation at the 3' *rny*_ends not corresponding to PNPase trimming stop positions (n=84) and at the 3' *rny*_ends corresponding to PNPase trimming stop positions (n=46). Upon RNase Y processing, PNPase degrades the RNA up to a putative secondary structure. **B.** Structure conservation at the total number of PNPase trimming stop positions

(n=183), which were previously identified in our laboratory[204]. The structure analyses were conducted by T.T.R. with A.L.R, A.-L.L. and L.B.

Next, we analysed the ΔG upstream and downstream of the total PNPase trimming stop positions, which were previously identified by our laboratory in *S. pyogenes*[204], using RNAfold (Figure 22B and see Material & Methods). We observed a decrease in ΔG just upstream of the total PNPase trimming stop positions identified, implying the presence of stable RNA structure (Figure 22B). It is possible that such structure could have a role in blocking PNPase degradation as already described for other bacteria *in vitro*[99,374,376]. In the previous study from our laboratory in *S. pyogenes*, the analysis of the average ΔG upstream of the total PNPase trimming stop positions did not reveal the presence of stable structures[204], in contrast to the current observation. This can be explained by the different window size parameter used in this analysis (see Material & Methods). In the aforementioned study, structures corresponding to terminator stem-loops were identified using a window size of 25 nt[204]. Here, larger and weaker structures than terminator stem-loops were identified by using a larger window size (50 nt) compared the one previously used.

In some cases, the 3' *rny*_ends, which corresponded to a PNPase trimming stop position, also correspond to an YhaM trimming stop position, as observed in the *SPy_0316* ORF (coding for a putative transcriptional regulator) and in the intergenic region between the Spy_sRNA73113 and *rplO* (coding for the 50S ribosomal protein L15) (Figure 23A and B). PNPase started degrading the Spy_sRNA73113 and *rplO* target RNAs at a G, corresponding to the initial RNase Y processing positions, located 34 or 120 nt upstream of the predicted trimming stop position, respectively. (Figure 23C and D). PNPase degraded the transcript until stem-loop structures which were predicted in both the targets, downstream of the trimming stop positions (Figure 23E and F). As suggested by the global prediction of RNA structures upstream of the total PNPase trimming stop position (Figure 22B), these RNA stem-loops could prevent PNPase degradative activity.

We also observed that YhaM trimmed these targets and while the trimming stop positions coincided with the ones of PNPase, the YhaM trimming start position was not detectable in the Δ*pnpA* strain, indicating the YhaM activity on these targets required PNPase. Thus, YhaM likely trimmed the RNA 3' end that was generated by the preceding PNPase trimming. This result indicates that YhaM also trims RNA 3' ends produced by other exoRNases besides the RNA 3' ends deriving from endoRNase processing or transcription termination. In the case of Spy_sRNA73113 and *rplO* intergenic region, the YhaM trimming stop position coincides with the previously identified 3' end of the Spy_sRNA7311[359], for which a terminator was not predicted. It is possible that the subsequent trimming of PNPase and YhaM, until the predicted stem-loop structure, is involved in the generation of the Spy_sRNA7311 3' end.

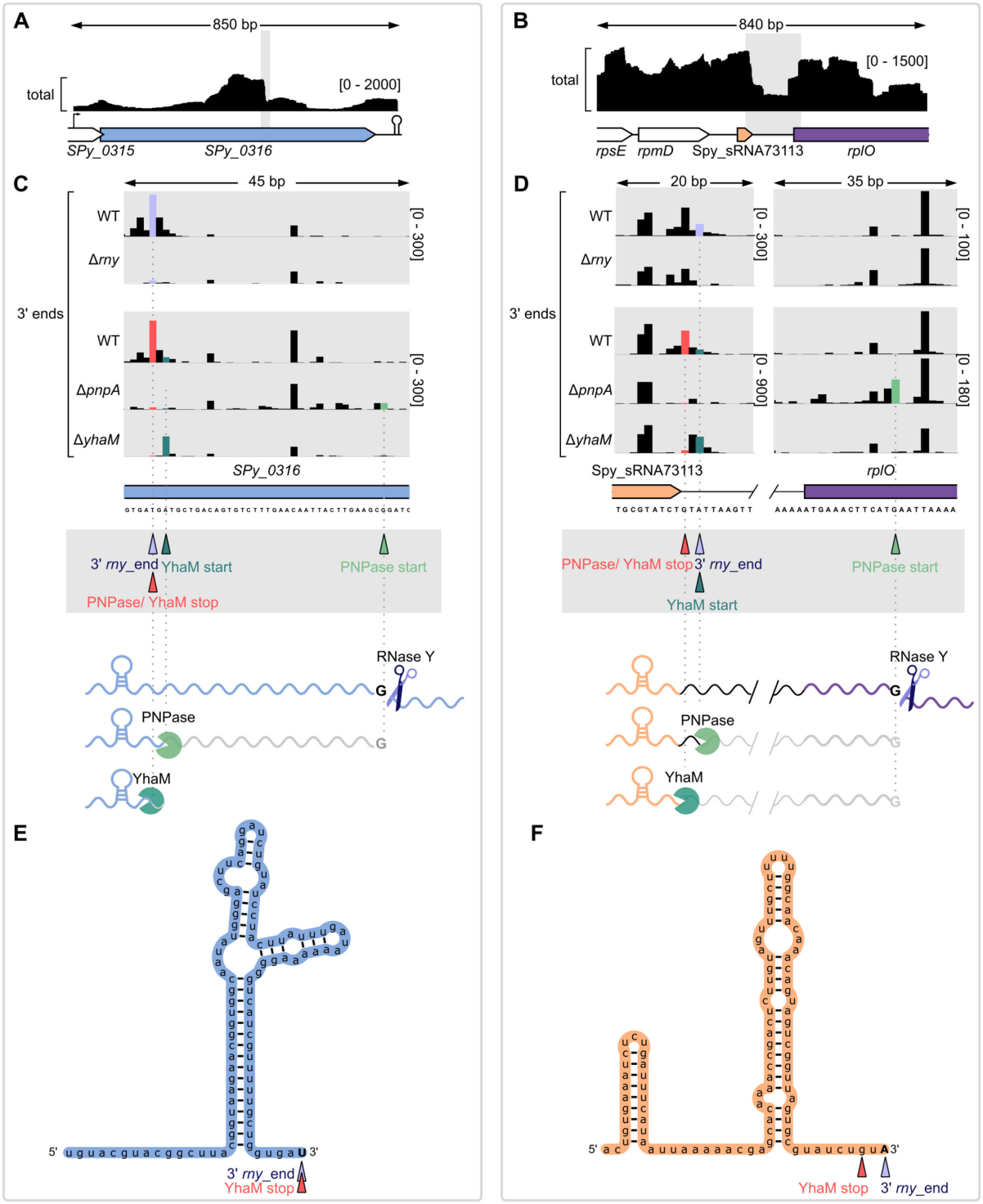


Figure 23. **Subsequent trimming of PNPase and YhaM after RNase Y processing.**
A and **B.** Total coverage profiles in the WT strain obtained from RNA sequencing and schematic representation of genomic loci: *SPy_0316* encoding a putative transcriptional regulator (in panel A) and the intergenic region between the putative sRNA Spy_sRNA73113 and *rplO* encoding the 50S ribosomal protein L15 (Spy_sRNA73113 ↔ *rplO*) (in panel B). The grey boxes highlight the regions where the processing positions of RNase Y, PNPase and YhaM were identified. **C** and **D.** On the top, 3' coverage profiles of *SPy_0316* and Spy_sRNA73113 ↔ *rplO* in the WT, *rny* (RNase Y) deletion mutant (Δ*rny*), *pnpA* (PNPase) deletion mutant (Δ*pnpA*) and *yhaM* (YhaM) deletion mutant (Δ*yhaM*) strains. The coverage scale is indicated between brackets. On the bottom, processing sites detected (3' *rny*_ends, trimming start positions, trimming stop positions depicted with purple, green and red arrowheads, respectively) are

indicated together with a schematic representation of the PNPase and YhaM ("pacman" symbols) trimming upon RNase Y (scissors) processing. For both the examples, RNase Y processed the transcripts at a guanosine (G), which corresponded to the location where PNPase started digesting these targets. PNPase stopped the trimming after 34 nt and 120 nt in *SPy_0316* and Spy_sRNA73113 ↔ *rplO*, respectively. The RNA 3' ends generated by PNPase were then targeted by YhaM that subsequently trimmed few nucleotides. **E** and **F.** RNA folding of the sequence 100 nt upstream of the 3' *rny*_ends corresponding to YhaM trimming stop position.

3.3.5 Short RNA fragments are produced by RNase Y and eventually trimmed by PNPase and/or YhaM

Several 3' *rny*_ends, targeted by PNPase and/or YhaM, corresponded to the RNA termini of fragments generated by RNase Y (Table A8). These RNA fragments were identified by comparing the nucleotide position of the 5' *rny*_ends and 3' *rny*_ends (Figure 24A, B and Table A8).

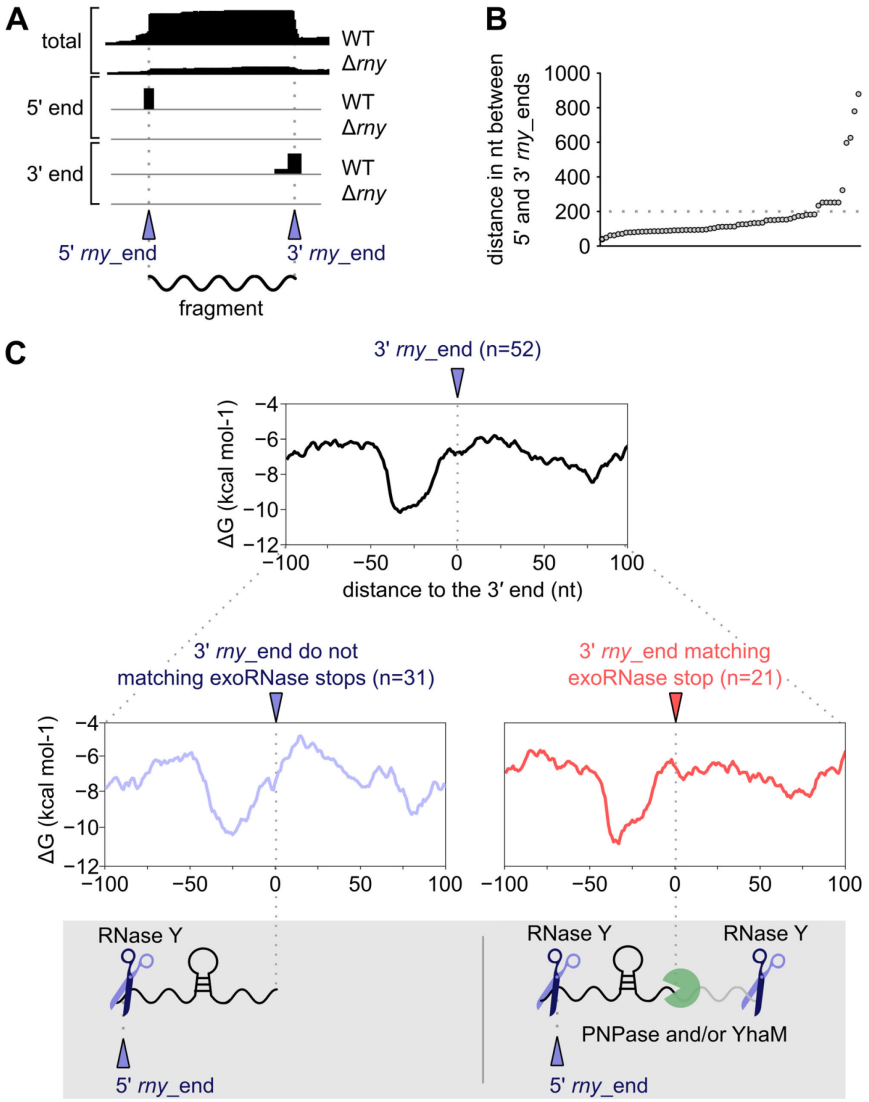


Figure 24. **Short fragments are produced by RNase Y and trimmed by PNPase and YhaM.**
A. Representation of total, 3' and 5' end coverage profiles from RNA sequencing in the WT and *rny* (RNase Y) deletion mutant (Δ*rny*) strains of short RNA fragments whose termini (5' *rny*_end and 3' *rny*_end)

production depends on RNase Y (Table A8). **B**. Comparison of 5' *rny*_end and 3' *rny*_end location, with a minimum and maximum distance set at 40 and 1000 nt, respectively. The paired 5' *rny*_ends and 3' *rny*_ends are depicted as dots. The analysis of the RNA fragments produced by RNase Y was performed by L.B. and A.-L.L. **C**. Structure conservation analysis at the total transcript fragment 3' ends (n=52, top), at the transcript fragment 3' ends that did not correspond to any 3'-to-5' exoRNase trimming stop positions (n=31, left) and at the transcript fragment 3' ends targeted by 3'-to-5' exoRNases (n=21, right). The minimum free energy (ΔG in kcal mol^{-1}) was calculated 100 nt upstream and downstream of the positions of interest with a sliding window of 50 nt. The structure analyses were conducted by T.T.R. with A.L.R, A.-L.L. and L.B. Below the ΔG graphs, schematic representation of the fragment production and subsequent trimming by 3'-to-5' exoRNase ("pacman" symbol). RNase Y (scissors) generates the transcript fragment 3' and 5' end. In some cases, PNPase and/or YhaM further trim the fragment 3' end.

The paired 5' *rny*_ends and 3' *rny*_ends were mainly found at 50 to 200 nt of distance, when allowing for a maximal separation of 1000 nt between the two *rny*_ends (Table A8). We observed that 60% of the fragment 3' ends were subjected to 3'-to-5' exoRNase trimming by PNPase and/or YhaM (Table A8). It is possible that the remaining 40% of fragment 3' ends were (i) trimmed but, the 3'-to-5' exoRNases trimming positions were not identified in our analysis, (ii) trimmed by other exoRNases (iii) not trimmed at all. In all the cases, by analysing the ΔG around the fragment 3' ends we observed a drop in ΔG upstream of the 3' *rny*_ends, indicating that the 3' fragment region was structured (Figure 24C). During RNA decay, the endoRNase-generated RNA fragments are usually rapidly degraded by exoRNases and undetectable in a WT strain. Here, the RNA fragments produced by RNase Y were instead present in the WT strain. The reason why these fragments were not completely degraded is unclear; they could correspond to decay intermediate fragments, that were not efficiently degraded because of the RNA structure at the fragment 3' ends (Figure 24C). Alternatively, these fragments could exert a biological role and therefore they were stabilized because necessary in the bacterial cell.

RNA fragments produced by RNase Y are specifically degraded in the absence of YhaM

Among the fragments described above, the ones deriving from the *murC* (encoding the UDP-N-acetylmuramate-L-alanine ligase), *SPy_1551* (encoding a hypothetical protein) and *SPy_0316* (encoding a putative transcriptional regulator) mRNAs were validated by Northern blot analysis (Figure 25). These fragments were detected in the WT strain, but not in the Δ*rny* strain, demonstrating that their production was RNase Y-dependent. These fragments were either trimmed by YhaM (Figure 25A and B) or by both YhaM and PNPase (Figure 25C). The exoRNase trimming start positions corresponded to the initial RNase Y processing position, which could be identified for the fragments derived from *murC* and *SPy_0316* mRNAs (Figure 25A and C; Table A8), as was previously described (see Figure 20B). Since YhaM trims on average only 3 nt at the

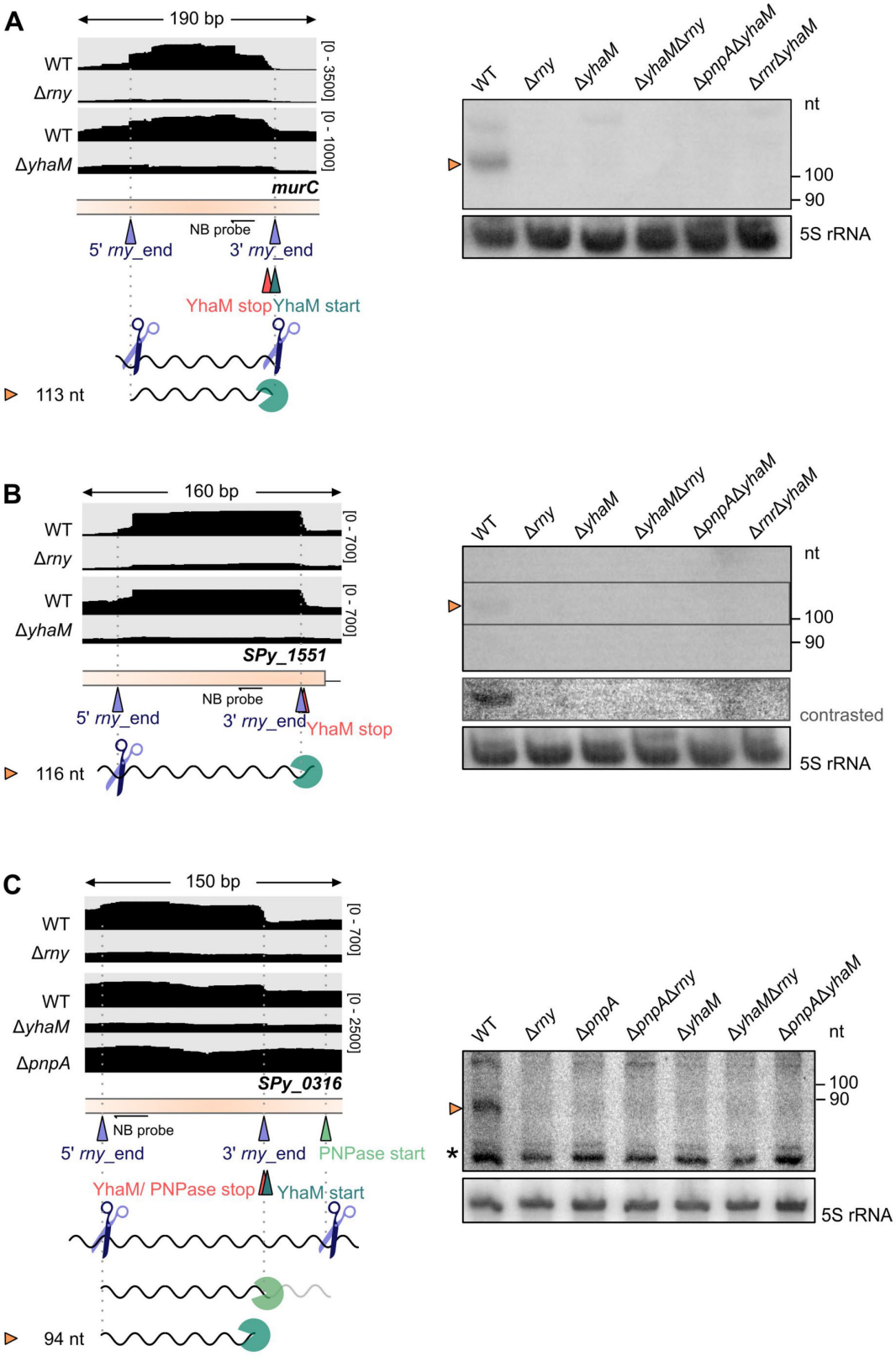


Figure 25. **RNase Y-generated RNA fragments are degraded in the absence of YhaM.**
Subset of fragments generated by RNase Y and trimmed by 3'-to-5' exoRNases. **A-C**. On the left, 3' coverage profile of portions of *murC* (encoding the UDP-N-acetylmuramate-L-alanine ligase), *SPy_1551* (encoding a hypothetical protein) and *SPy_0316* (encoding a putative transcriptional regulator) in the WT, *rny* (RNase Y) deletion mutant (Δ*rny*), *yhaM* (YhaM) deletion mutant (Δ*yhaM*) strains and also in the *pnpA* (PNPase)

deletion mutant (Δ*pnpA*) strain in panel C. The coverage scale is shown between brackets. The 5' and 3' *rny*_ends (purple arrowheads), the 3'-to-5' exoRNase start and stop positions (green and red arrowheads) are indicated. RNase Y (scissors) generates both the fragment 5' and 3' ends. The fragment 3' ends were trimmed by YhaM or by PNPase/YhaM. On the right, Northern blot analyses of the fragments in *murC*, *SPy_1551* and *SPy_0316*. In the WT, Δ*rny*, Δ*yhaM*, *yhaM* and *rny* double deletion mutant (Δ*yhaM*Δ*rny*), *pnpA* and *yhaM* double deletion mutant (Δ*pnpA*Δ*yhaM*) and *rnr* (RNase R) and *yhaM* double deletion mutant (Δ*rnr*Δ*yhaM*) strains. For the fragment in *SPy_0316*, the Northern blot analysis was performed also in the Δ*pnpA* and in the *rny* and *pnpA* double deletion mutant (Δ*rny*Δ*pnpA*) strains. The labelled oligonucleotide probes used for the Northern blot analyses (NB probe) are shown below the loci representation and coverage profiles. When the fragments were poorly detectable a contrasted portion of the blot is shown. The 5S rRNA was used as a loading control. The results of one Northern blot analysis (n=3) are shown. In panel C, the star depicts an RNA not specifically recognized by the probe. K.H. contributed in generating the replicates of the Northern blot analyses.

RNA 3' ends, we at first expected to detect fragments that were few nucleotides longer in the Δ*yhaM* strain than in the WT strain. Conversely, the fragments were not detected in the Δ*yhaM* strain by both RNA sequencing and Northern blot analyses (Figure 25A-C). We next predicted that these fragments were degraded by PNPase and/or RNase R in the absence of YhaM. To test this hypothesis the Northern blot analyses were also performed in the double deletion mutant Δ*pnpA*Δ*yhaM* and Δ*rnr*Δ*yhaM* strains (Figure 25A-C). However, the fragments were also not detected in these strains (Figure 25A-C), suggesting that either RNase R and PNPase acted redundantly or that another RNase was involved in the fragment degradation.

3.3.6 PNPase completely degrades the RNAs generated by RNase Y processing

In the comparative analysis of the 3' *rny*_ends with the 3'-to-5' exoRNase trimming start and stop positions, we observed that most of the RNA 3' ends were further trimmed upon RNase Y processing. The 5' *rny*_ends were in 87.4 % of the cases located at G and in only 12.6% of the cases not mapped at a G. The RNA 5' ends, starting at a G, likely coincides with the RNase Y processing positions, while the other could result from the subsequent activity of other RNase(s) (Figure 26A). This observation indicates that while most of the RNA 3' ends produced by RNase Y were trimmed by 3'-to-5' exoRNases, the RNA 5' ends were not targeted by other RNases.

As already mentioned before, after RNase Y processing either one or both the RNase Y-generated RNA products were degraded (*i.e.* either the 3' *rny*_end or the 5' *rny*_end was detected) (Figure 18A). Therefore, the RNA 3' ends produced during the same processing events responsible for the generation of the 190 5' *rny*_end could not be identified (Figure 26A). To examine the fate of the undetected RNA products upstream of the 190 5' *rny*_ends, the last were paired to the 3'-to-5' exoRNase trimming start positions that were located up to 10 nt before the 5' *rny*_ends (Figure

26B and Table A9; see Material & Methods). We observed that 12% of the 5' *rny*_ends were in vicinity of the PNPase trimming start positions (Figure 26B and C).

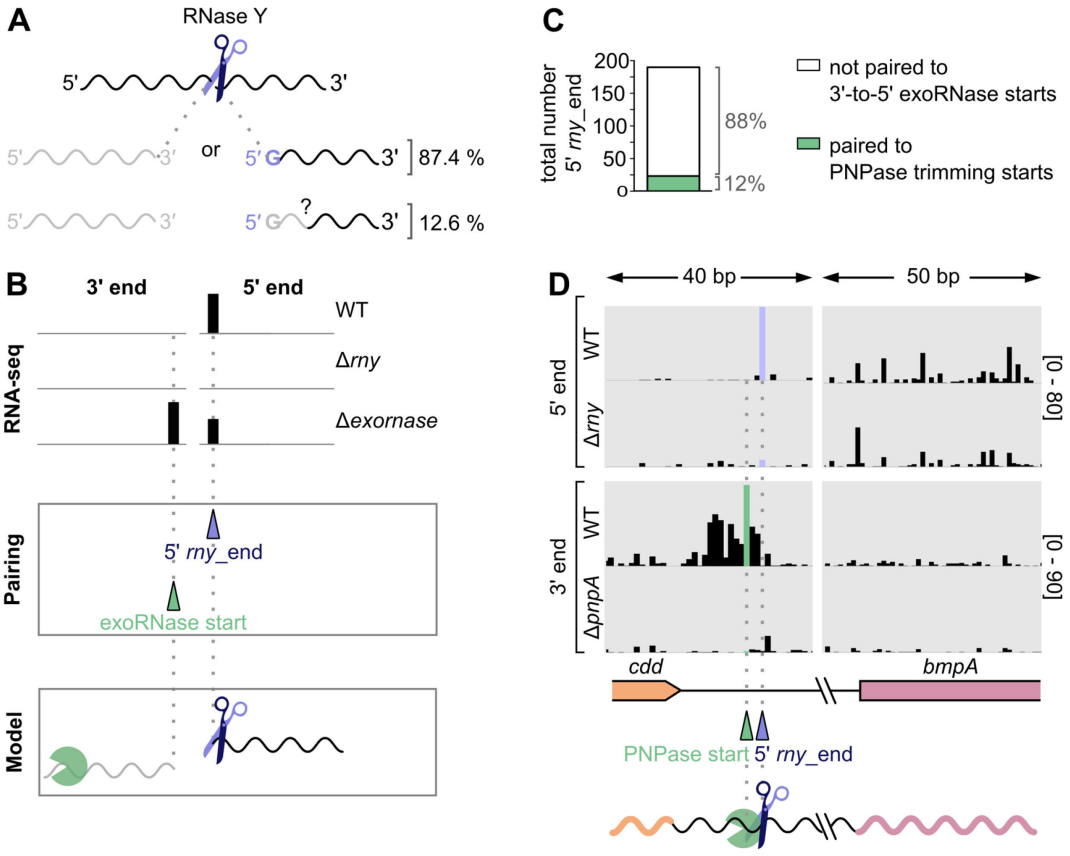


Figure 26. **PNPase degrades completely the transcripts processed by RNase Y.**
A. Representation of RNase Y (scissors) processing event generating an RNA 5' and 3' end. The two newly generated RNA ends were never simultaneously detected. The RNA 5' ends detected were in 87.4% located at a G, likely at the initial RNase Y processing site, and in 12.6% of the cases were not found upstream of a G suggesting that they were further processed/trimmed after the RNase Y processing event. **B.** Identification of the RNA products upstream of the RNase Y processing positions (*i.e.* upstream of the detected 5' *rny*_ends). The 5' *rny*_ends were compared with the 3'-to-5' exoRNase trimming start positions located at least 10 nt upstream, (see Material & Methods and Table A9). The pairing of the 5' *rny*_ends with PNPase trimming start positions was performed by T.T.R. with A.L.R, A.-L.L. and L.B. RNA-seq: schematic representation of the 3' and 5' end coverage profiles from RNA sequencing; Comparison: the RNA ends that were included in the comparative analysis are indicated. The RNA ends more abundant in the WT strain than in the *rny* (RNase Y) deletion mutant (Δ*rny*) strain and the 3'-to-5' exoRNase trimming start positions are depicted with purple and green arrowheads, respectively; Model: After RNase Y processing, PNPase completely degrades the RNA product upstream of the RNase Y processing site. **C.** The bar plot shows the portion of RNA 5' ends (5' *rny*_ends) that were paired to 3'-to-5' exoRNase trimming start positions (bottom in green), and the portion of the 5' *rny*_ends which were not associated to any 3'-to-5' exoRNase trimming start (top in white). All the 3'-to-5' exoRNase trimming start positions paired to 5' *rny*_ends were PNPase trimming starts. **D.** 5' and 3' end coverage profiles from RNA sequencing of portions of the *cdd* (coding for a cytidine deaminase) and *bmpA* (coding for a lipoprotein) ORFs and the intergenic region between the two genes in the WT, Δ*rny* and *pnpA* (PNPase) deletion mutant strains. The coverage scale is indicated between brackets. The 5' *rny*_end (purple arrowheads) and the PNPase trimming start position (green arrowheads) are indicated. RNase Y (scissors) processed the *cdd-bmpA* intergenic region subsequent targeted by PNPase ("pacman" symbol).

This result suggested that PNPase was responsible for the complete degradation of the RNA products upstream of the RNase Y processing event (Figure 26C and Table A9). The trimming start positions of PNPase were not identified exactly one nucleotide upstream of 5' *rny*_ends, but these positions were in most of the cases up to 4 nt apart (Table A9). This divergence is likely due to the involvement of YhaM, which in general trims few nucleotides from the RNA 3' ends. An example of a 5' *rny*_end paired with a PNPase trimming start position was identified in the intergenic region between *cdd* (coding for a cytidine deaminase) and *bmpA* (coding for a lipoprotein) (Figure 26D). The 5' *rny*_end, previously validated by primer extension analysis (Figure 11C), was found 3 nt upstream of a PNPase start position. Since the PNPase trimming stop position was not identified, we concluded that PNPase was most likely responsible for degrading completely the RNA containing the *cdd* ORF.

The majority of the 5' *rny*_ends (88%) were not paired to any PNPase trimming start position (Figure 26C), therefore the degradation mechanism of the RNA products upstream of the RNase Y processing positions was not deciphered.

3.3.7 Conclusions III

We characterized the RNase Y-dependent RNA 3' ends and observed that they exhibit different characteristics from the RNA 5' ends (*e.g.* lack of G conservation). Therefore, the origin of the RNase Y-dependent RNA 3' ends (3' *rny*_ends) was further investigated by using an RNA sequencing comparative approach. In most of the cases the 3' *rny*_ends resulted from trimming of 3'-to-5' exoRNases. In particular, we concluded that PNPase is the main 3'-to-5' exoRNase that acts in concert with RNase Y in RNA degradation.

3.4 Investigating the role of RNase Y and PNPase interplay

3.4.1 RNase Y and PNPase allow differential RNA stability of the *rsmC-cdd* and *bmpA* RNAs

Based on the previous observation that RNase Y and PNPase targeted the *cdd-bmpA* intergenic region (Figure 26), we further explored the role of both RNases on the post-transcriptional regulation of the *cdd* and *bmpA* mRNAs. The *cdd* and *bmpA* genes are part of a longer operon including *rsmC* (coding for a 16 rRNA methyltransferase), *SPy_1227*, *SPy_1226* and *SPy_1225* (coding for a sugar ABC transporter) (Figure 27A). To explore the effect of RNase Y and PNPase activity in the *cdd-bmpA* intergenic region, rifampicin assays followed by Northern blot analyses were performed with labelled oligonucleotide probes targeting *cdd* (Figure 27B) and the *bmpA* mRNAs (Figure 27C). Transcription driven from the *rsmC* promoter to the predicted transcriptional terminator downstream of the *bmpA* ORF led to the production of a ~ 2900 nt long primary transcript that was highly stabilized in the Δ*rny* strain compared to the WT strain (Figure 27A and B). The increased stability of the primary transcript in the Δ*rny* strain is explained by the fact that RNase Y cleaved this transcript in the UTR between *cdd* and *bmpA*, as previously demonstrated (Figure 26). RNase Y activity generated a ~ 1700 nt RNA product upstream of the processing site, which comprises the *rsmC* and *cdd* ORFs (Figure 27A and B). This RNA isoform was scarcely detectable in the WT strain and stabilized in the Δ*pnpA* strain, indicating that it was likely degraded by PNPase (Figure 27B). This observation suggested that RNase Y processing provided access to PNPase, which subsequently degraded the generated *rsmC-cdd* RNA isoform. Contrary to the latter isoform, the stability of the ~ 1200 nt *bmpA* RNA isoform, coinciding with the RNA product downstream of the RNase Y processing site, did not vary between the WT and Δ*pnpA* strains and its half-life was higher than the one of the *rsmC-cdd* RNA isoform (Figure 27A and C). These results provided evidence that RNase Y and PNPase activity ensured the differential decay of the *rsmC-cdd* and *bmpA* transcripts.

The primary *rsmC-cdd-bmpA-SPy_1227-SPy_1226-SPy_1225* polycistronic transcript is also processed by RNase III in the intergenic region between *bmpA* and *SPy_1227*[185]. Two RNA 5' ends produced by RNase III were annotated; one was identified bioinformatically (in blue) and the other by manual screening of the RNA sequencing data (in light blue) (Figure 27D). RNase III produced these two RNA 5' ends by processing both sides of a stem loop, generating the characteristic 2 nt 3' overhang in the dsRNA products[179] (Figure 27D). Deletion of *rnc* (encoding RNase III) led to the accumulation of a long transcript (> 6000 nt) corresponding to the full operon transcript isoform that was not detectable in the WT strain (Figure 27E).

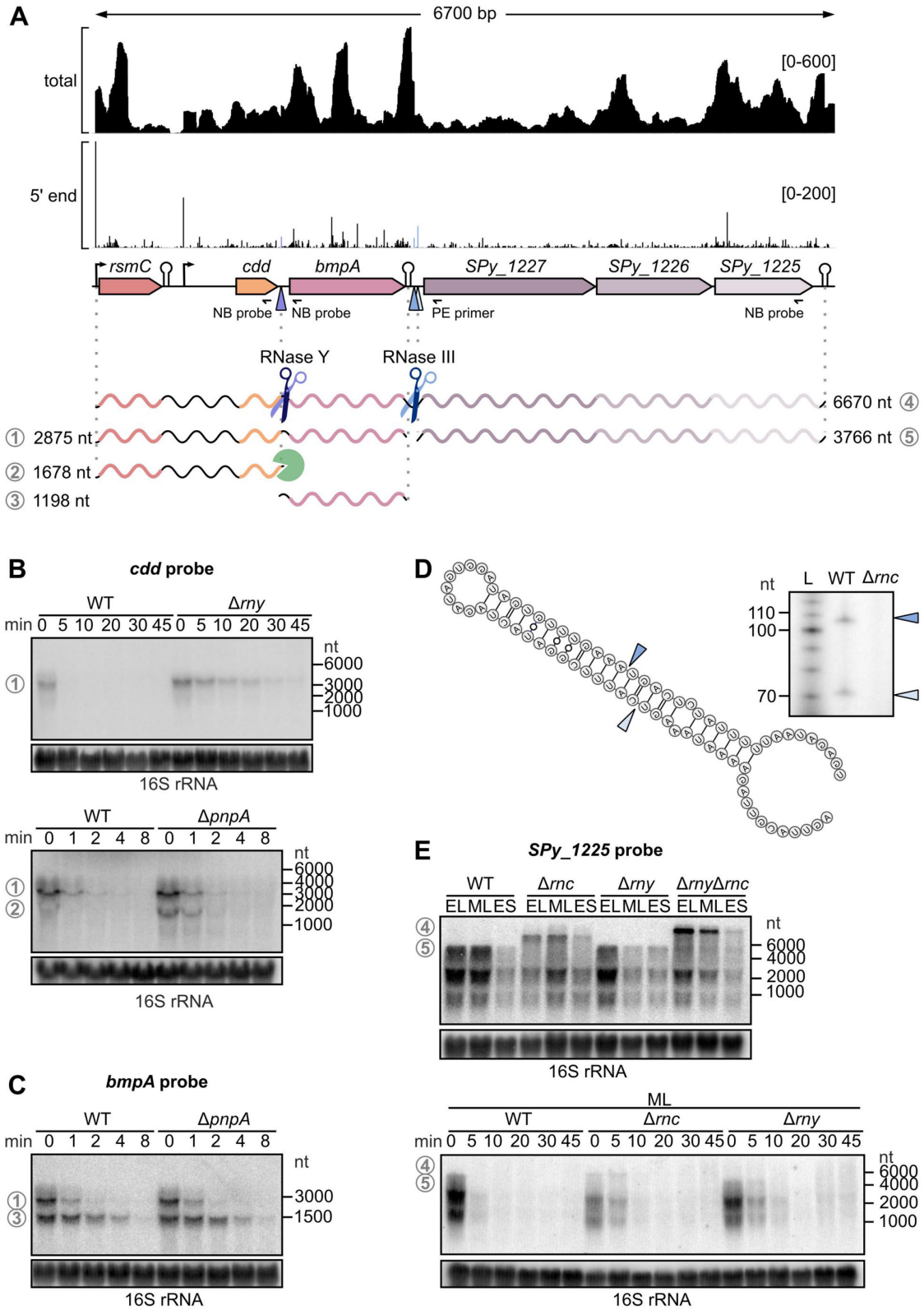


Figure 27. **RNase Y and PNPase allow differential stability of co-transcribed mRNAs.**
A. Total and 5' end coverage profiles from RNA sequencing of the *rsmC-cdd-bmpA-SPy_1227-SPy_1226-SPy-1225* operon in the WT strain. The scales are indicated between brackets. Representation of the operon with the predicted promoters, terminators, RNase Y processing site (purple arrowhead) and RNase III

processing sites (blue arrowheads) that we previously identified[185]. The labelled oligonucleotide probes used in the Northern blot analyses (NB probe) and the primer used in the primer extension analysis (PE primer), are depicted with black arrows below the respective targeted locus. The RNA isoforms detectable by Northern blot analyses and processed by RNase Y (purple scissors), by RNase III (blue scissors) and eventually degraded by PNPase ("pacman" symbol) are represented with the expected sizes in nt. **B.** and **C.** Transcript stability study by Northern blot analyses, up to 45 or 8 minutes (min) after the addition of the rifampicin in the WT and in the *rny* (RNase Y) deletion mutant (Δ*rny*) strains or in the *pnpA* (PNPase) deletion mutant (Δ*pnpA*) strain. The 16S rRNA was used as a loading control. **D.** On the left: RNA folding of the region surrounding the RNase III processing sites (20 nt upstream and downstream of the two processing sites, respectively) in the untranslated region between *bmpA* and *SPy_1227*[185]. On the right: primer extension analyses in the WT and *rnc* (RNase III) deletion mutant (Δ*rnc*) strains at mid-logarithmic growth phase. The primer extension analysis was performed with A.-L.L. **E.** The operon *rsmC-cdd-bmpA-SPy_1227-SPy_1226-SPy-1225* expression profile was assessed by Northern blot analysis at early-logarithmic (EL), mid-logarithmic phase (ML) and early-stationary (ES) growth phases in the WT, Δ*rnc*, Δ*rny* and *rny* and *rnc* double deletion mutant (Δ*rny*Δ*rnc*) strains. The transcript stability was analysed at mid-logarithmic phase of growth in the WT, Δ*rny* and Δ*rnc* strains. The 16S rRNA was used as a loading control. The results of one representative Northern blot analysis (n=3) are shown.

Therefore, it is likely that RNase III cleaved the readthrough RNAs after the *rcmC-cdd-bmpA* terminator, separating the *SPy_1227-SPy_1226-SPy_1225* RNA isoform from the *rsmC-cdd-bmpA* RNA isoform, which was cleaved by RNase Y.

3.4.2 RNase Y produces short decay intermediate fragments rapidly degraded by PNPase

As described in the introduction, RNA degradation is usually initiated by an endoRNase processing that leads to the production of decay intermediate fragments normally degraded by exoRNases and therefore not detectable in the WT strain. It was shown in *S. pyogenes* that PNPase is involved in the degradation of 185 decay intermediate fragments of 50 to 200 nt in length[204]. The sequence around these fragment 3' and 5' ends was analysed (Figure 28A). Strikingly, we observed the conservation of a G just upstream of the decay intermediate fragment 5' ends (*i.e.* RNA ends more abundant in the Δ*pnpA* strain than in the WT strain and named 5' Δ*pnpA*_ends) (Figure 28A). Therefore, we deduced that RNase Y was responsible for the production of most of these decay intermediate fragment 5' ends. To support this hypothesis, two decay intermediate fragments (produced from *pyrH* and *SPy_2197* mRNAs) were examined by Northern blot analyses in the WT, Δ*rny*, Δ*pnpA* and Δ*rny*Δ*pnpA* strains (Figure 28B and C). The decay intermediate fragments were detectable in the Δ*pnpA* strain but not in the Δ*pnpA*Δ*rny* strain (Figure 28B and C). This result confirms that RNase Y participates in the generation of the decay intermediate fragment 5' ends. We did not observe any G conservation at the decay intermediate fragment 3' ends suggesting that they could originate from the processing of an unidentified endoRNase or from RNase Y cleavage and subsequent trimming by a 3'-to-5' exoRNase.

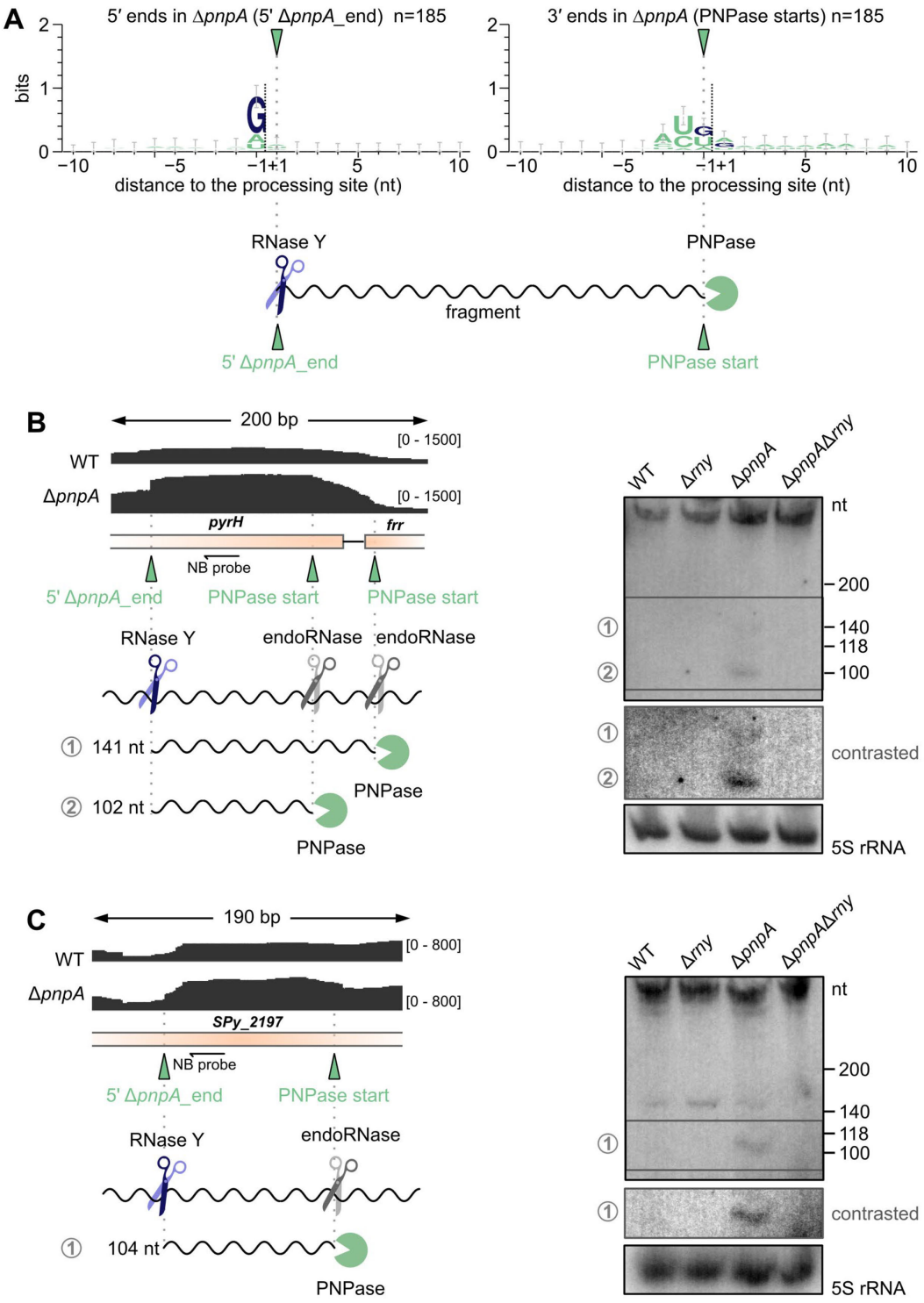


Figure 28. **PNPase degrades decay intermediate fragments generated by RNase Y.**
A. Alignment of the sequences around the RNA 5' and 3' ends of decay intermediate fragments (n=185) that were previously identified only in the Δ*pnpA* strain[204]. The fragment 5' end, which is more abundant in the Δ*pnpA* strain than in the WT strain, is indicated as 5' Δ*pnpA*_end. Below the sequence alignments (conducted with T.T.R, A.L.R. and A.-L.L.) a schematic representation of the fragment 5' and 3' ends generation and degradation is depicted. **B** and **C**. On the left, total coverage profile from RNA sequencing of the decay intermediate fragments identified in *pyrH-frr* and in *SPy_2197*, the coverage scales are reported

between brackets. The 5' Δ*pnpA*_end and the PNPase trimming start positions are indicated with green arrowheads. RNase Y (purple scissors) and an unidentified endoRNase (grey scissors) generate a decay intermediate fragment subsequently degraded by PNPase. The labelled oligonucleotide probe (NB probe) used in the Northern blot analyses and the predicted size of the fragments are shown. On the right, Northern blot analyses in the WT, *rny* (RNase Y) deletion mutant (Δ*rny*), *pnpA* (PNPase) deletion mutant (Δ*pnpA*) and in the *pnpA* and *rny* double deletion mutant (Δ*rny*Δ*pnpA*) strains at mid-logarithmic growth phase. When the fragment was poorly detectable, a contrasted region of the blot is shown. The 5S rRNA was used as a loading control. The results of one representative Northern blot analysis (n=3) are shown. Northern blot analyses were performed together with K.H.

3.4.3 Decay of 5' regulatory elements depends on the concerted action of RNase Y and PNPase

We further analysed the decay intermediate fragments produced by RNase Y and degraded by PNPase and we observed that they often derived from the processing of regulatory RNA 5' UTRs (*e.g.* T-boxes and riboswitches) (Figure 29). A role for PNPase in the degradation of decay intermediate fragments deriving from 5' regulatory elements was previously reported by our laboratory[204]. (Figure 29A). In the thiamine pyrophosphate riboswitch (TPP) a 3' *rny*_end coinciding with a PNPase trimming start position was identified (Figure 29A and Table A7). This RNA 3' end corresponded to the 3' termini of a decay intermediate fragment that was already identified in the TPP riboswitch[204]. This fragment was detectable in the Δ*pnpA* strain, indicating that it was degraded by PNPase as previously shown[204], but was not present in the Δ*rny*Δ*pnpA* strain, suggesting that the fragment generation was depending on RNase Y (Figure 29A). Therefore, the decay of the TPP riboswitch occurred through the endoribonucleolytic processing of RNase Y, which generated an RNA 3' end targeted by PNPase. The fragment 5' end, which was originally identified as a 5' end more abundant in the Δ*pnpA* strain than in the WT strain (5' Δ*pnpA*_end), was likely generated by an unidentified endoRNase.

RNase Y also processed the *serS*, *SPy_1570_SPy_1569_valS* and *thrS* 5' UTRs, generating the RNA 3' end of the decay intermediate fragment, whose 5' end coincided with the regulatory element TSS (Figure 29B-D). We did not identify any PNPase trimming start position in the 5 nt upstream of the 3' *rny*_ends in the T-boxes *serS* and *valS* (Table A7 and see Material & Methods). However, based on the Northern blot analysis results we concluded that the decay intermediate fragments originating from *serS* and *valS* 5' UTRs were degraded by PNPase (Figure 29B and C). To pinpoint the PNPase trimming start positions, we extended the search to 1000 nt upstream of the 3' *rny*_ends (see Material & Methods) and we identified a PNPase trimming start position that coincides with RNase Y processing site in both *serS* and *valS* (Figure 29B and C). Indeed, the length of the *serS* and *valS* decay intermediate fragments detectable by Northern blot analysis in the

Δ*rny*Δ*pnpA* strain corresponded to the distance between PNPase trimming start positions (*i.e.* RNase Y processing sites) and the TSS of the regulatory element.

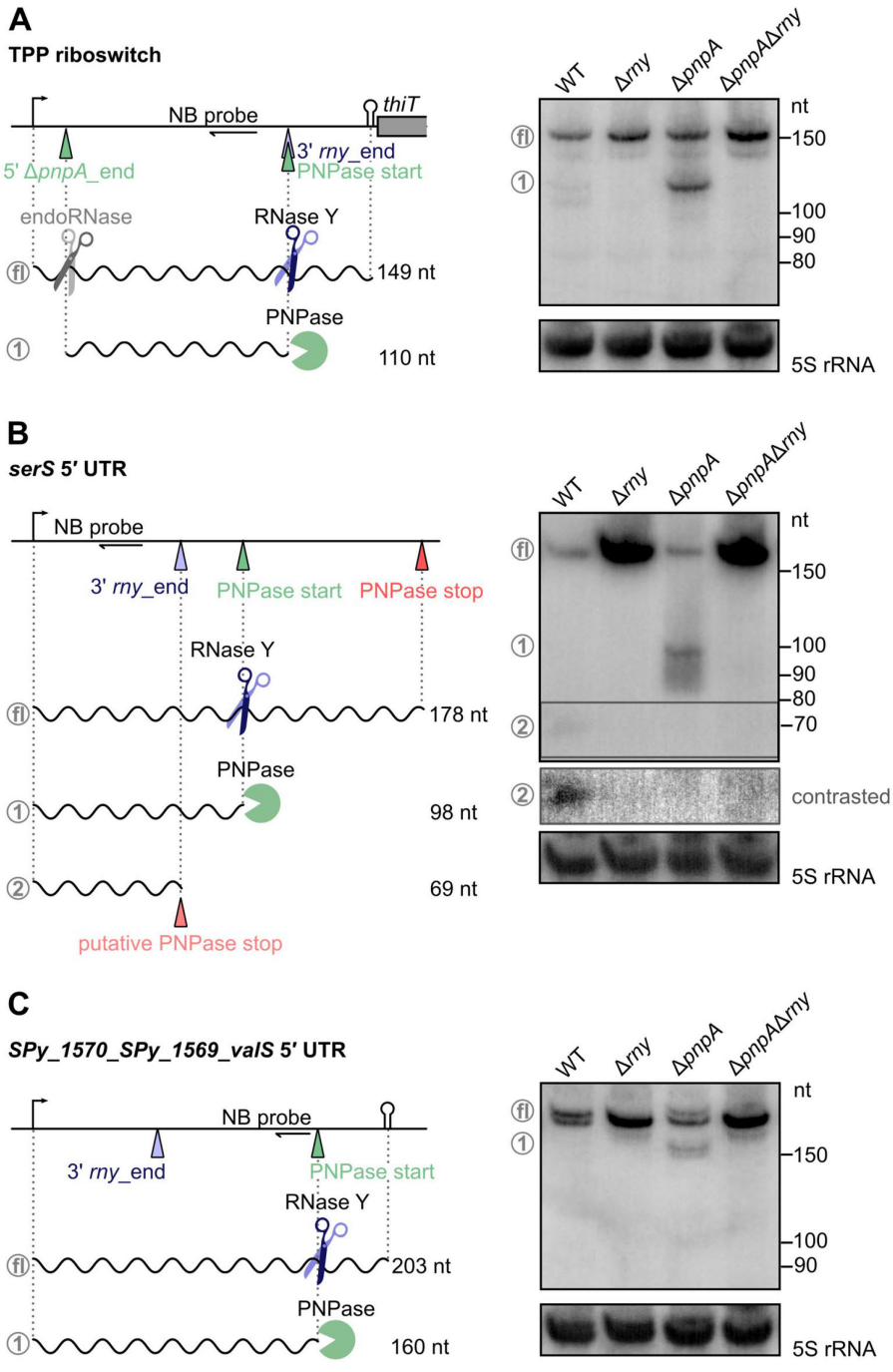
In the *serS* 5' UTR, a short fragment was visible only in the WT strain and the fragment 5' and 3' end corresponded to the *serS* TSS and to the 3' *rny*_end, respectively (Figure 29B). It is likely that during the degradation of the *serS* decay intermediate fragment, PNPase paused at the 3' *rny*_end (putative PNPase stop position), thereby producing this short fragment isoform (Figure 28B).

The decay of the *serS* and *valS* T-boxes was also shown to be initiated by RNase Y in *S. aureus*[160]. The authors observed that the RNA products downstream of the RNase Y processing site were detectable only in a strain expressing an inactivated variant of the 5'-to-3' exoRNase J1, suggesting that this exoRNase was responsible for the degradation of these fragments[160]. While the decay intermediate fragment in *serS* was also shortened at the 3' end by 3'-to-5' exoRNase, the one in *valS* was degraded by RNase J1 up to the transcriptional terminator[160]. Here we did not investigate the fate of the RNase Y-generated downstream fragments in *serS* and *valS* T-boxes, but we studied the degradation of the fragment upstream of the RNase Y processing site and observed that PNPase was responsible for the decay of these products. We do not exclude that RNase J1 is also involved in the *serS* and *valS* T-box decay, as described for *S. aureus*.

In the *thrS* 5' UTR, we identified two PNPase trimming start positions that did not correspond to any *rny*_ends. By Northern blot analysis, we deduced that PNPase was responsible for the degradation of two decay intermediate fragments deriving from *thrS* 5' UTR[204] (Figure 29D). To examine the origin of these fragments, we performed Northern blot analyses also in the Δ*pnpA*Δ*rny* strain and we observed that they were present in the Δ*pnpA* strain, but not in the Δ*pnpA*Δ*rny* strain, demonstrating that their production was RNase Y-dependent (Figure 29D). These are examples of how the screening for endoRNase processing sites in an exoRNase deletion mutant facilitates the identification of the endoRNase processing events.

Finally, only for the *glyQ* T-box RNase R was also involved in the degradation of the RNase Y-generated decay intermediate fragment, beside PNPase (Figure 29E). However, while PNPase stopped the degradation of this decay intermediate fragment downstream of the TSS[204], RNase R completed it (Figure 29E). An alternative processing position was detected in the absence of RNase Y (*i.e.* RNA end more abundant in the Δ*rny* strain than in the WT strain and named 5' Δ*rny*_end) (Figure 29E, in grey and Table A10). Therefore, it is likely that another endoRNase, by substituting RNase Y, is responsible for the degradation of the *glyQ* T-box. Indeed, the Northern blot analysis revealed the presence of a short RNA fragment detectable only in the Δ*rny* stain (Figure 29E).

Overall, these results point to the conclusion that the interplay of RNase Y and PNPase is responsible for the decay of the transcripts resulting from premature termination of T-box and riboswitch leaders.



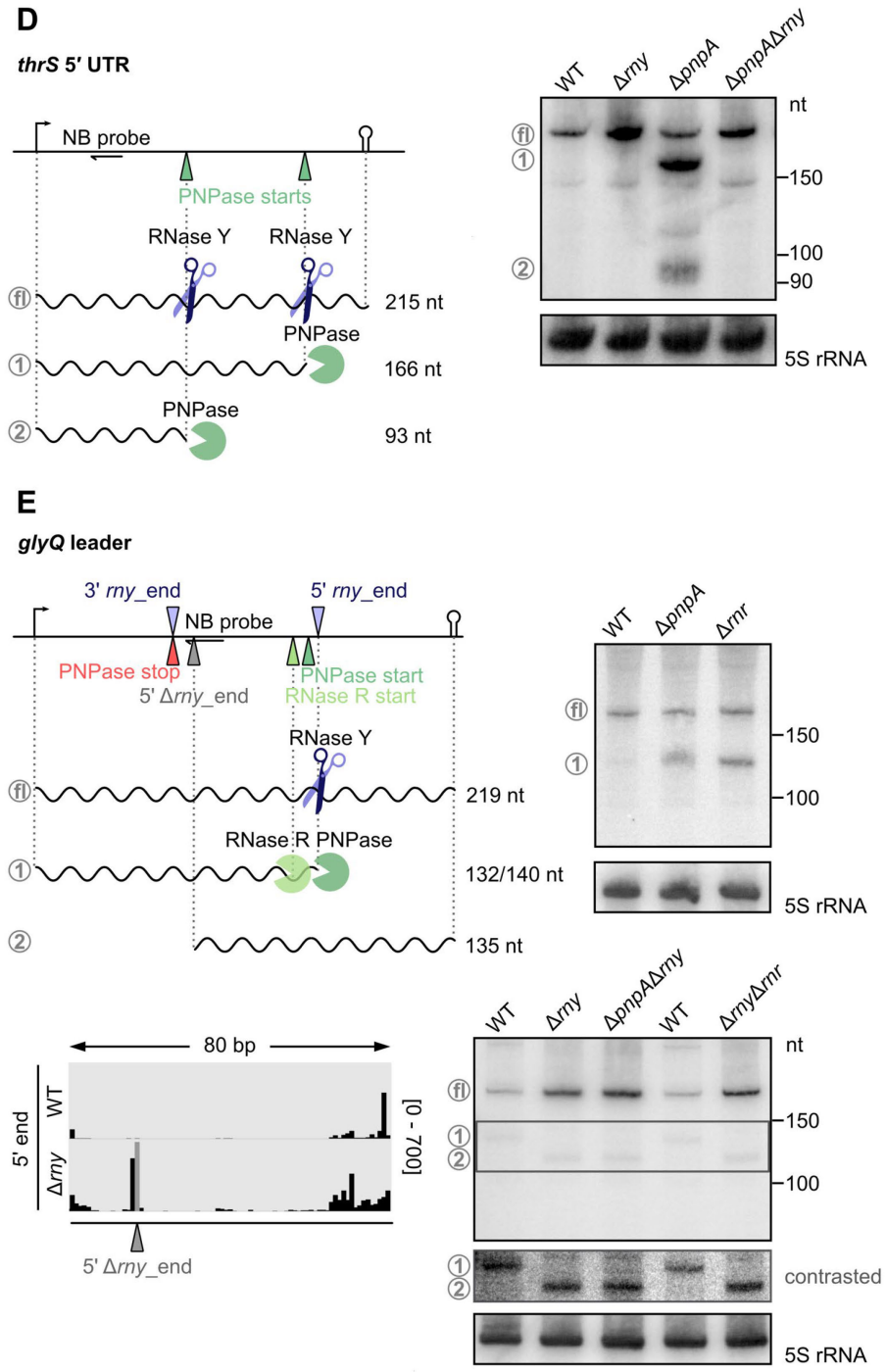


Figure 29. **Role of RNase Y and PNPase in the degradation of 5' regulatory elements.**
A-E. Representation of the predicted regulatory elements (*i.e.* T-boxes or riboswitches) with the predicted promoter, terminator, 5' *rny*_ends (purple arrowheads), 3'-to-5' exoRNase trimming start and stop positions (green and red arrowheads, respectively). The labelled oligonucleotide probes used in the Northern blot analyses and the expected size of the full length (fl) regulatory element transcripts and of the decay intermediate fragments, generated by RNase Y (purple scissors) and degraded by PNPase and/or RNase R ("pacman" symbols), are indicated. Northern blot analyses of the 5' regulatory element in the WT, *rny* (RNase Y) deletion mutant (Δ*rny*), *pnpA* (PNPase) deletion mutant (Δ*pnpA*), *rny* and *pnpA* double deletion mutant (Δ*rny*Δ*pnpA*), *rnr* (RNase R) deletion mutant (Δ*rnr*) strains and also in the *rny* and *rnr* double deletion mutant (Δ*rny*Δ*rnr*) strain for *glyQ* T-box. The 5S rRNA was used as a loading control. The results of one representative Northern blot analysis (n=3) are shown. **E.** In the *glyQ* T-box an RNA 5' end more abundant

in the Δ*rny* strain than in the WT strain was identified and named 5' Δ*rny*_end (grey arrowhead). Below the 5' regulatory element representation, the 5' end coverage profiles in the WT and Δ*rny* strains, obtained from RNA sequencing, is shown. The coverage scale is indicated between brackets. K. H. generated the replicates of the Northern blot analyses of the 5' regulatory elements.

3.4.4 RNase Y affects the *rpsB* and *rpsB-tsf* transcript stability

The abundance of the *rpsB* (encoding the 30S ribosomal protein S2) and *tsf* (encoding the elongation factor thermos-stable, EF-Ts) transcripts was strongly increased in the Δ*rny* strain compared to the WT strain (Table A1). In addition, proteomic analysis performed in both the WT and Δ*rny* strains indicated that the EF-Ts protein was overexpressed in the Δ*rny* strain (data not shown). To decipher whether this upregulation was due to an increase in transcript stability, we performed a stability assay followed by Northern blot analyses (Figure 30A-C). By probing *rpsB* and *tsf* mRNAs, we detected two and one transcript isoforms, respectively (Figure 30A-C). The longer RNA isoform corresponded to the full-length transcript, comprising both *rpsB* and *tsf* ORFs. Transcription driven from the *rpsB* promoter up to the predicted terminator in the *rpsB-tsf* intergenic region gave rise to the second transcript isoform, corresponding to the *rpsB* monocistronic transcript (Figure 30A). All the transcript isoforms were strongly stabilized in the Δ*rny* strain (half-life greater than 45 minutes) compared to the WT strain (half-life shorter than 5 min) (Figure 30B and C).

RNase Y was previously shown to affect the expression of *rpsB-tsf* operon also in *S. aureus*[160]. In this bacterium the *rpsB-tsf* operon architecture differs for the one observed in *S. pyogenes*, as *tsf* harbours its own promoter and a small ORF is located in the *rpsB-tsf* intergenic region[160]. In the abovementioned study, RNase Y processing sites were identified in both *rpsB* mRNA and *rpsB-tsf* UTR. However, we did not detect any RNase Y processing events in this operon. We instead observed, in the *rpsB* 5' UTR, several 5' RNA ends present in the Δ*rny* strain, but not in the WT strain (5' Δ*rny*_end) (Table A10), suggesting that another RNase is acting on the *rpsB* transcript in the absence of RNase Y (Figure 30D and E). In addition, two PNPase trimming stop positions were previously identified in the *rpsB* 5' UTR, by our laboratory[204] (Figure 30D). The abundance of the RNA 3' ends corresponding to these PNPase trimming stop positions was reduced in the Δ*rny* strain compared to the WT strain, suggesting that PNPase activity in the *rpsB* 5' UTR was depending on RNase Y. Two short RNA isoforms were detected by Northern blot analyses in the WT strain, which could originate from the *rpsB* TSS and terminate at the two PNPase trimming stop positions (Figure 30D and F). These two RNA isoforms, detectable up to 8 min after the addition of the rifampicin, were not present in the Δ*pnpA* and Δ*rny* strains, indicating their production requires both RNase Y and PNPase activity. Different RNA isoforms were detected in

the absence of RNase Y (Δ*rny* and Δ*rny*Δ*pnpA*) and likely originated from the alternative processing events identified in the Δ*rny* strain (Figure 30D and F).

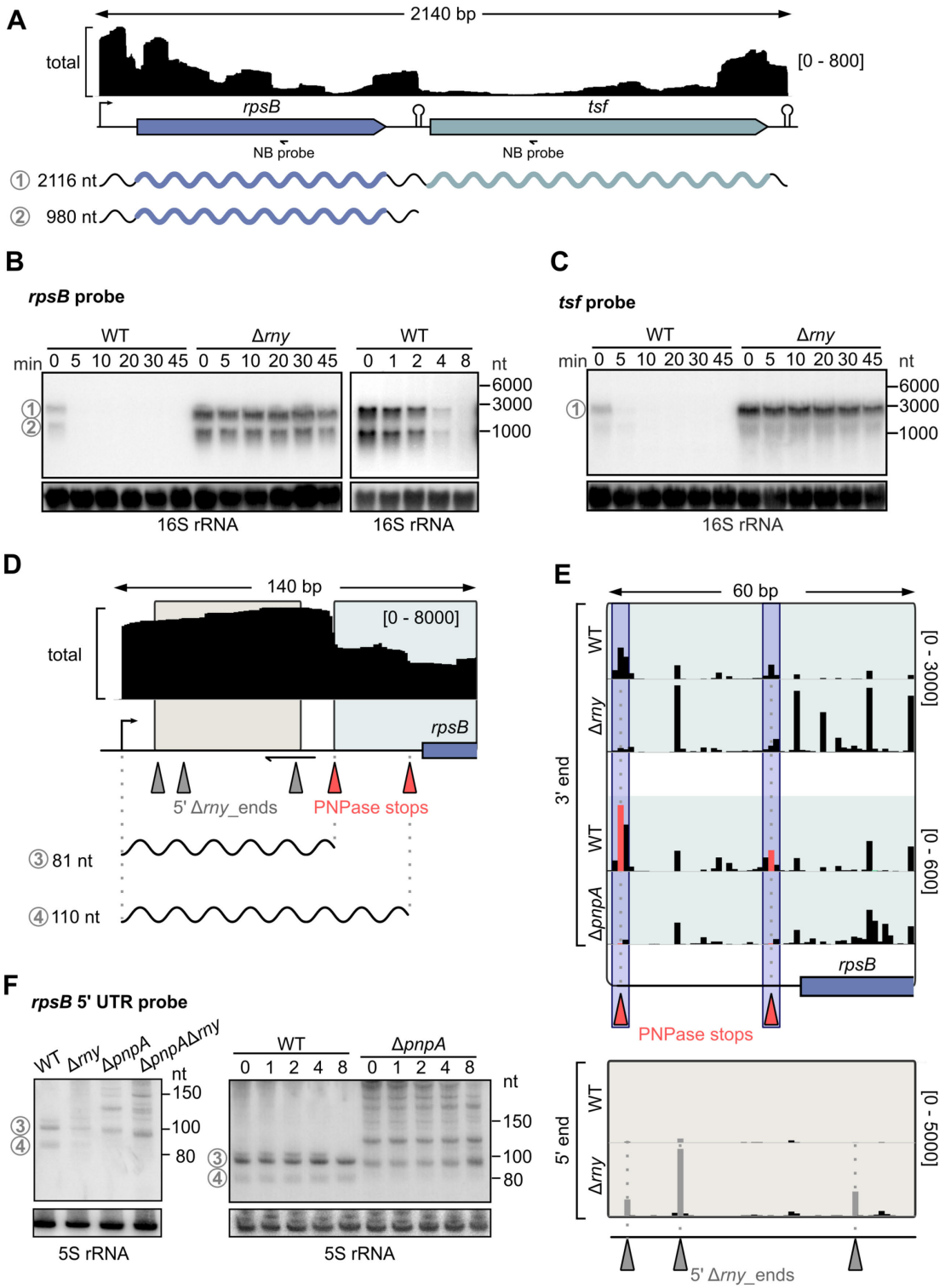


Figure 30. **Stabilization of the *rpsB* and *rpsB-tsf* transcripts in the absence of RNase Y.**
A. Total coverage profile of the *rpsB-tsf* operon in the WT strain, obtained from RNA sequencing. The scale is indicated between brackets. Below the coverage profile, the *rpsB-tsf* operon is schematically represented with the predicted promoter, terminators and the labelled oligonucleotide probes used for the Northern blot analyses (NB probes). The transcripts detectable in Northern blot analyses are shown with the expected

sizes. **B.** and **C.** Transcript stability studied by Northern blot analyses, up to 45 or 8 minutes (min) after the addition of the rifampicin in the WT and in *rny* (RNase Y) deletion mutant (Δ*rny*) strains. The 16S rRNA was used as a loading control. The results of one representative Northern blot analysis (n>3) are shown **D.** Total coverage profile from RNA sequencing of the *rpsB* 5' UTR and a portion of *rpsB* ORF with the identified PNPase trimming stop positions (red arrowheads) and the RNA 5' ends more abundant in the Δ*rny* strain than in the WT strain (5' Δ*rny*_ends, grey arrowheads). The rectangles highlight the portions of the *rpsB* 5' UTR further analysed in panel E. **E.** 3' end coverage profiles in the WT, Δ*rny* and *pnpA* (PNPase) deletion mutant (Δ*pnpA*) strains and 5' end coverage profiles in the WT and Δ*rny* strains obtained from RNA sequencing. The coverage scales are indicated between brackets. The trimming stop positions within the purple box are likely to be RNase Y dependent. **F.** On the left, Northern blot analysis with a labelled oligonucleotide probe targeting the *rpsB* mRNA 5' UTR (see primer in panel D) in the WT, Δ*rny*, Δ*pnpA* and *rny* and *pnpA* double deletion (Δ*rny*Δ*pnpA*) strains. On the right, transcript stability was evaluated by Northern blot analysis up to 8 minutes (min) after the addition of the rifampicin in the WT and Δ*pnpA* strains. The 5S rRNA was used as a loading control. The results of one representative Northern blot analysis (n=3) are shown. The study in the *rpsB* mRNA 5' UTR by Northern blot analysis was conducted with A.-L.L. and K.H.

Overall, it is unclear how the *rpsB* and *rpsB-tsf* transcript stability is strongly affected in the absence of RNase Y in *S. pyogenes*. Here, we showed that both RNase Y and PNPase participate in the production of short RNA isoforms deriving from the *rpsB* mRNA 5' UTR. Of note, a putative sRNA of 136 nt deriving from the *rpsB* mRNA 5' UTR was previously annotated, based on the structure conservation of this 5' UTR in *E coli*[26,377]. Successively, in this bacterium, the *rpsB* mRNA 5' UTR was shown to play a regulatory role in *cis*, as it is recognized by the 30S ribosomal protein S2 (encoded by *rpsB*), which in turn affects its own and *tsf* expression[378]. In the possibility that the S2 protein autoregulates its own expression at post-transcriptional level also in *S. pyogenes*, it is possible the RNase Y and PNPase activity in the *rpsB* mRNA 5' UTR is involved in this process.

3.4.5 Conclusions IV

We investigated the interplay between RNase Y and PNPase and we observed that these two enzymes are involved in the RNA degradation of 5' regulatory elements and in uncoupling the stability of mRNAs that are part of the same polycistronic transcript. Importantly, the study of RNase Y activity in the Δ*pnpA* strain, by the transcriptome comparison, permitted the identification of processing events not detectable otherwise. We therefore provide a more accurate annotation of the RNase Y targetome in *S. pyogenes*. Overall, from the study of the RNase Y and PNPase concerted action we demonstrated that our RNA sequencing comparative approach successfully allows to study the bacterial RNase interplay and mechanisms of RNA degradation *in vivo*.

Discussion

4.1.1 Role of RNase Y in the regulation of *speB* expression

In this thesis, to gain insight into the broad activity of RNase Y in *S. pyogenes* we first determined the genes that were differentially expressed in the Δ*rny* strain compared to the WT strain. We identified 80 genes that were either upregulated or downregulated in the absence of RNase Y. Considering that M1 *S. pyogenes* strain SF370 has a genome with 1801 genes annotated[363,379], in the Δ*rny* strain only 4.4% of the transcriptome is affected, at mid-logarithmic growth phase and in rich medium (Figure 7 and Table A1). A previous study revealed that in the absence of RNase Y, the expression of almost 30% of the transcriptome was altered. However, this analysis was performed in M14 *S. pyogenes* H5C5 strain, at stationary growth phase and using a medium poor in carbohydrates and rich in amino acids[168]. Therefore, it is possible that the role of RNase Y in transcript degradation might differ depending on the environmental conditions, including nutrients availability.

Different functional gene categories were affected in the absence of RNase Y, comprising genes involved in bacterial metabolism and in protein synthesis (Table A1). The expression of several virulence genes (*slo*, *speB*, *grab* and *sic*) was altered in the Δ*rny* strain. This observation was consistent with the known role of RNase Y in virulence genes regulation described in *S. aureus*[167] and *C. perfringens*[166] and also in *S. pyogenes*[168]. The *slo* gene, encoding the Streptolysin O, and *speB*, encoding for the SpeB protease, were already shown to be regulated by RNase Y[168,329]. However, while *slo* was described to be upregulated in the Δ*rny* strain, here we observed that *slo* mRNA abundance was decreased in the Δ*rny* strain. This observation can be explained by the different strain and growth conditions used in this thesis compared to the previous study[168]. We identified also two virulence genes (*i.e. grab* and *sic*) that were not previously reported as regulated by RNase Y. Whereas *grab* encodes for a G-related alpha2-macroglobulin-binding protein, a surface protein that inhibits extracellular proteases[380,381], the *sic* gene encodes the Streptococcal inhibitor of complement, which protects the bacteria from lysis caused by the complement system.

It was shown that RNase Y is required for *speB* expression in *S. pyogenes*[168,329], and our data confirmed this observation (Figure 8). The SpeB protease is one of the major virulence factors associated to *S. pyogenes* pathogenesis and toxicity, therefore it is of interest to understand how the production of this enzyme is regulated. We therefore aimed at deciphering the role of RNase Y in the regulation of *speB* expression and we reported that (i) RNase Y indirectly affects *speB* transcription (Figure 9) and (ii) RNase Y processes the *speB* mRNA 5' UTR (Figure 13) producing a stable putative sRNA (Figure 17).

4.1.2 RNase Y impact on *speB* transcription

The study of the *speB* promoter activity revealed that, in the absence of RNase Y, the transcription from *speB* promoter was strongly reduced (Figure 9). From this observation we deduced that *speB* expression is controlled at the transcriptional level by RNase Y. In an effort to pinpoint the unknown intermediate factor involved in the RNase Y-mediated regulation of *speB* mRNA abundance, we tested the impact of RNase Y on the expression of known *speB* transcriptional regulators. RopB (*speB* activator) and the CovRS two-component system (*speB* repressor) are the best characterized *speB* direct transcriptional regulators and previous observation indicated that RNase Y is involved in their regulation[106], therefore they were selected for further analysis (Figure 10C and D). Here we observed that both *ropB* and *covRS* transcript stability was increased in the Δ*rny* strain. If this increased transcript stability would correlate with a higher amount of RopB and CovR protein levels, then CovR (but not RopB), could be a possible factor involved in the downregulation of *speB* in the Δ*rny* strain. This hypothesis could be validated by generating a double deletion mutant of *rny* and *covRS* or by mutating the CovR binging site in the *speB* promoter region and subsequently examining the *speB* promoter activity. Thus, the impact of RNase Y on the protein levels of these regulators should be investigated. It has been recently shown that RopB production is reduced in the Δ*rny* strain compared to the WT strain, although the *ropB* mRNA abundance was identical in the two strains[371]. In this case, the downregulation of *speB* expression in the Δ*rny* strain could be explained by the decrease of RopB levels. However, it must be noted that in the abovementioned study, the *S. pyogenes* NZ131 serotype M49 was used and in this strain the *ropB* transcript pattern differs from the one observed in the strain SF370 serotype M1 studied in this thesis[106]. It is also possible that in the absence of RNase Y, RopB production varies among different strains.

As described in the introduction, *speB* expression is controlled by several regulators and likely additional unknown factors. Therefore, it is likely that the reduction of *speB* expression in the Δ*rny* strain results from the action of several regulators, which are controlled by RNase Y. For

instance, the SIP peptide, which indirectly induces *speB* expression by enhancing RopB activity[315], is co-transcribed with *speB* (Figure 8A and B) and its expression is downregulated in the Δ*rny* strain. A SIP-based negative feedback signal could possibly cause *speB* downregulation in this strain.

Overall, while we demonstrated that the reduction of *speB* abundance in the Δ*rny* strain is due to an indirect effect on transcription, a better characterization of the regulatory pathway(s) affected by RNase Y is needed. This may also shed light on the impact of RNase Y on other virulence factors (*sic*, *slo* and *grab*), which were found differentially expressed in the Δ*rny* strain (Table A1). In *S. aureus*, promoter activity analysis of virulence genes revealed that RNase Y regulates their expression indirectly at the transcriptional level[167]. The authors observed that the RNase Y-mediated regulation of virulence genes did not depend on the major *S. aureus* virulence regulators (*i.e.* the RNAIII sRNA and the two-component system SaeRS), but it was likely due to yet unidentified factors[167]. The authors hypothesized that, among these unknown regulatory factors, several sRNAs, which were differentially expressed in the Δ*rny* strain, could play a critical role in the RNase Y control of virulence gene expression. In our study we identified six putative sRNAs whose expression was altered in the Δ*rny* strain (Table A1) and we do not exclude their possible involvement in the RNase Y-mediated regulation of virulence genes, including *speB*.

4.1.3 The effect of RNase Y processing in *speB* mRNA 5' UTR

We observed that RNase Y also cleaves the *speB* transcript in the 5' UTR region (Figure 13). The effect of these processing events on *speB* expression and/or SpeB production is still unclear and further analyses should aim at dissecting the impact of these cleavage sites on *speB* post-transcriptional regulation. In *S. aureus*, RNase III can enhance translation by shortening the 5' UTR of mRNAs[184] and we questioned whether RNase Y could similarly affect SpeB production. Therefore, we analysed the expression of a reporter gene fused to *speB* mRNA 5' UTR and expressed under the control of a constitutive promoter, which is not controlled by RNase Y (data not shown). The results from this translational fusion experiment were inconclusive, as the generated measurements were not reproducible. In addition, it is worth noting that the *speB* 5' UTR is also processed by other unidentified RNase(s) (Figure 13), thus the impact of also this processing event on *speB* mRNA translation should be assessed.

Recently it was demonstrated that the truncation of *speB* mRNA 5' UTR alters the temporal pattern of *speB* expression and causes a general increase of the bulk mRNA stability in *S. pyogenes*[161]. These findings indicate that the *speB* mRNA 5' UTR, could display regulatory functions. In the present study, we observed that the RNase Y processing of the *speB* mRNA 5' UTR results in the production of a short RNA fragments (Figure 17), which could play regulatory roles, for instance,

as a *trans*-acting sRNA. A putative sRNA (Spy_sRNA1699993) was already annotated in the *speB* 5' UTR[359] and we demonstrated that the sRNA 5' end coincides with the RNase Y cleavage site(s) (Figure 17). Although the hypothesis of a regulatory sRNA deriving from *speB* 5' UTR (named here *speB*-sRNA) requires validation, several observations sustain this possibility: *speB*-sRNA is (i) produced at a specific growth phase, (ii) generated by endoRNase processing, which is a common biogenesis strategy for sRNAs[29,30,53,154], (iii) stable over time, and (iv) further processed by unidentified RNase(s), which could play a role either in the sRNA maturation or possibly in the sRNA function.

4.2 RNase Y targetome and processing determinants

The differential expression analysis, performed by comparing the RNA abundance in the WT and Δ*rny* strains, provides a global overview of the transcripts affected by RNase Y, but it does not allow to distinguish between direct and indirect targets. To gain a better understanding on RNase Y activity, we next aimed at pinpointing the direct targets of RNase Y (*i.e.* RNAs cleaved by RNase Y), by mapping the RNA 5' and 3' ends in both the WT and Δ*rny* strains (Figures 11 and 18). Mapping of the RNA ends more abundant in the WT strain than in the Δ*rny* strain revealed 190 and 130 RNA 5' and 3' ends that derived from RNase Y processing, respectively. While most of the identified RNA 3' ends resulted from trimming of 3'-to-5' exoRNases upon RNase Y processing (Figure 20), the RNA 5' ends corresponded to the original RNase Y processing position (Figures 11 and 26). The different origin of the RNA 5' and 3' ends will be further discussed in the next section 4.3.

Some of the RNA 5' and 3' ends were annotated within the same target and the distance between these ends ranged between 50 to 200 nt (Figure 24A and B, Table A8). However, since the RNA 3' ends do not correspond to the original RNase Y processing positions, but rather to the trimming stop positions of 3'-to-5' exoRNases, it is likely the average distance between two RNase Y processing positions is longer than observed.

The analysis of RNA 5' ends – coinciding to the original RNase Y processing positions – allows to examine the target determinants for RNase Y processing. Consistently with the fact that RNase Y is a single-stranded endoRNase, we found that the minimum free energy increased at the RNA 5' ends produced by RNase Y (Figure 11D). Among the RNA 5' ends identified, 87.4% were located downstream of a G (Figure 11D). An RNase Y preference to cleave the transcripts downstream of a G was already identified in *S. aureus*, where 58% of the RNA 5' ends produced by RNase Y were annotated after this nucleotide[160]. The apparent difference in the frequency of G at the RNase Y processing sites in the two bacteria can be explained, for example, by: (i) the diversity

in the RNA sequencing workflow and data analysis and (ii) the role of RNase J1, which might be more active in *S. aureus* than in *S. pyogenes* in degrading the RNA 5' ends produced by RNase Y and (iii) the difference in the RNase Y orthologue activity.

As the two RNase Y cleavage sites identified in *speB* mRNA 5' UTR were mapped after a G (Figure 13), we used this mRNA 5' UTR as a model target to investigate the importance of the G for RNase Y activity. We provided *in vivo* evidence that the G located downstream of the RNase Y cleavage sites is essential for RNase Y processing of the *speB* mRNA 5' UTR (Figure 15). Therefore, based on the genome-wide mapping of the RNase Y cleavage sites, in combination with the analysis of the RNase Y processing activity in the *speB* mRNA 5' UTR, we concluded that RNase Y requires a G immediately downstream of the processing position to cleave the target. Interestingly, when the G was mutated in the *speB* mRNA 5' UTR (*i.e.* the G was substituted in A), alternative processing events were detected (Figure 15). From the observation that these alternative cleavage sites were also mapped at a G, we hypothesized that they resulted from RNase Y activity. If this was the case, RNase Y would be still able to recognize the *speB* mRNA as a target, in the absence of the G, and other factors would be likely involved in the RNase Y identification of the cleavage site. It is possible indeed that, while the G could be important for the RNase Y catalytic activity, other determinants are necessary for the cleavage site recognition.

In light of the limited number of RNase Y cleavage sites identified in this study (320 5' and 3' RNA ends), it is surprising that only a G can limit RNase Y activity, since all the transcripts in the cell would contain multiple RNase Y cleavage sites. Despite the promiscuity of RNase Y, the processing of an RNA target depends also on the accessibility of this RNase to it. Indeed, the Gs within an RNA molecule could be often masked by ribosomes, interacting proteins and sRNAs that limit the number of G available for RNase Y recognition. This has been demonstrated to be the case for RNase E in *E. coli*, in which the linear scanning, for cleavage site search, within ssRNA regions is impeded by obstacles – including bound proteins, RNA duplexes and ribosome occupancy – that inhibits the processing downstream[115]. The ribosome-mediated protection from processing did not occur from shielding the cleavage site, but rather by the intense ribosome occupancy, which depends on the strength of the Shine-Dalgarno sequence[115].

The specificity for one single nucleotide is not unique to RNase Y, but it is employed by other RNases, including the fungal endoRNase T1, which also cleaves ssRNA after a G[382]. As a result, RNase T1 is largely used in molecular biology to cleave RNAs *in vitro* and to predict RNA secondary structures. RNase T1 preference for the G is explained at a structural level, as the guanosine is able to form several hydrogen bonds with conserved amino acids of the RNase catalytic site[383]. On the other hand, RNase E, from *S. enterica*, was demonstrated to cleave

preferentially 2 nt upstream of a U. The interaction of RNase E with this residue promotes a conformational change of the enzyme itself, which favours the catalytic activity and therefore the cleavage of the RNA[30]. The RNase Y crystal structure is not yet available and therefore it is unclear how the G is recognized and why it is important for the catalytic activity. The impact of the G on the RNase Y activity *in vitro* has not been investigated and RNase Y *in vitro* studies are in general limited[109]. For instance, recent attempts of studying RNase Y catalytic activity *in vitro* were performed using a great excess of protein over the RNA substrate, indicating that this enzyme is extremely inefficient in the conditions tested[109,384]. This could be due to: (i) the lack of co-factors or interacting proteins that are required for RNase Y activity in the reaction, (ii) the lack of interacting RNAs, for instance sRNAs and (iii) secondary or tertiary RNA structures that might not occur at the conditions used *in vitro*. Examples of these three possibilities are described below.

It has been recently shown that in *B. subtilis*, RNase Y interacts with three proteins YlbF, YmcA and Yaat. This complex influences RNase Y specificity, by favouring RNase Y processing of certain specific mRNAs[153]. Besides interacting proteins, sRNAs can also modulate RNase Y activity. This is the case of the VR-RNA sRNA that induces a structural change of the mRNA target and thereby it promotes RNase Y processing[166,232] (exemplified in Figure 2C). In addition, RNase Y activity was also shown to be affected by the presence of secondary structure a few nucleotides (6 nt) downstream of the processing position. This have been demonstrated for two targets: the *yitJ* S-adenosylmethionine riboswitch in *B. subtilis*[119] and the *saePQRS* transcript in *S. aureus*[169]. However, in our study, in agreement with previous transcriptomic analysis performed in both *B. subtilis* and *S. aureus*, we did not identify RNA secondary structures in proximity of the RNA 5' ends produced by RNase Y (Figure 11D). Analysis of RNA secondary structures in the *speB* mRNA 5' UTR also did not reveal the presence of a structure just after the RNase Y processing events (Figure 16). Importantly, bioinformatic prediction of complex RNA structures is still challenging at a genome-wide scale. Therefore, although it seems that RNase Y in *S. pyogenes* does not depend on the recognition of structured RNA, we do not exclude that, at least for specific targets, RNA structures might play a role in RNase Y specificity.

Overall, different requirements for RNase Y processing have been identified in the Gram-positive bacteria studied so far. It is possible that different RNase Y orthologs employ distinct strategies to recognize the targets and/or that various determinants affect in concert RNase Y activity.

4.3 The interplay of RNase Y and exoRNases

In order to study the role of RNase Y in RNA degradation we compared the RNase Y targetome, with the 3'-to-5' exoRNase (PNPase, YhaM or RNase R) trimming start and stop positions, which were previously characterized[204]. Although the notion that endoRNases act in concert with exoRNases to perform RNA degradation is well established, it has been validated only for a limited number of targets and in few bacterial species. Here, an RNA sequencing comparative approach was used to demonstrate the interplay of RNase Y with 3'-to-5' exoRNases, genome-wide at the nucleotide resolution.

We observed that 58% of the identified RNA 3' ends generated by RNase Y were trimmed mainly by PNPase and/or YhaM (Figure 20). In agreement with the minor role of RNase R in RNA degradation at standard growth conditions[204], we observed very little interplay between RNase Y and RNase R, with only one RNase Y-dependent RNA 3' end further trimmed by RNase R (Tables A5, A7, A9). Since most of the RNA 3' ends did not map at a G (Figure 18C), we hypothesized that also the remaining 42% of RNA 3' ends were targeted by exoRNases. Conversely, most of the RNA 5' ends (87.4%) corresponding to the initial RNase Y processing positions were mapped at a G (Figure 11D) and were not further trimmed. The remaining 12.6% of the RNA 5' ends was not located at a G and could derive either from trimming, likely performed by the 5'-to-3' exoRNase J1, or from processing events by an endoRNase, whose activity depends on RNase Y.

The method to identify the RNase Y processing sites, developed in our laboratory[204], depends on the detection of at least the RNA 5' or 3' end generated by processing event. Therefore, we were unable to identify RNase Y cleavage sites when both the RNA products upstream and downstream of the RNase Y processing site were degraded, for instance by 3'-to-5' exoRNases. Therefore, it is possible that the portion of RNAs detected as cleaved by RNase Y represents only a subset of the total RNase Y targets. This implies that (i) the number of RNase Y processing site, identified in this study, is likely an underestimation and that (ii) we cannot study the interplay of RNase Y and 3'-5' exoRNases in the degradation of the undetected RNAs (*i.e.* fully degraded RNAs). In addition, the method does not allow the detection of processing position that derived from the activity of redundant RNases (*i.e.* alternative RNase(s) complementing the activity of the missing RNase). Therefore, it is possible, that the 42% of RNA 3' ends not corresponding to any trimming positions in our analysis, could derive from redundant trimming of different 3'-to-5' exoRNases.

Considering the fate of only the detected RNA ends produced by RNase Y, it seems that the RNA 3' ends are subjected to exoRNase activity, while the RNA 5' ends are resistant to further

degradation. The different susceptibility to RNA degradation between RNA 5' and 3' ends suggests that the interplay between RNase Y and RNase J1 is limited compared to the one observed between RNase Y and PNPase. This is possibly due to (i) higher PNPase protein levels compared to RNase J1, (ii) higher RNA degradation rates for PNPase than RNase J1, (iii) both the abovementioned scenarios. Alternatively, the RNA product downstream of the RNase Y processing site could be released, from the catalytic site into the cytosol, slower than the upstream fragments, explaining why the newly generated RNA 5' end is protected from degradation.

4.3.1 RNase Y mainly acts in concert with PNPase in RNA degradation

It was previously shown that PNPase is the major 3'-to-5' exoRNase participating in RNA degradation in *S. pyogenes*[204] and *B. subtilis*[136]. It was therefore not surprising that PNPase appeared to be responsible for the trimming of most of the RNA 3' ends produced by RNase Y (Figure 20). The two enzymes have been shown to physically interact in *B. subtilis* and the amino acids involved in this interaction have also been elucidated[227]. However, an *in vivo* localization study revealed that PNPase does not co-localize at the membrane with RNase Y, indicating that the interaction of these enzymes is likely transient[229]. In *S. pyogenes*, it is not known whether RNase Y and PNPase are associated, as the sole protein that was demonstrated to interact with RNase Y so far is the enolase enzyme. It is possible that in *S. pyogenes*, PNPase and RNase Y are physically close to each other, explaining why the two enzymes appear to act rapidly in concert in RNA degradation.

The decay intermediate fragments produced by endoRNases are usually efficiently degraded by PNPase and become detectable only in the Δ*pnpA* mutant[123,136,204]. The sequence conservation analysis of the decay intermediate fragments identified in the Δ*pnpA* strain revealed the presence of a G at the RNA 5' ends of 127 decay intermediate fragments (Figure 28A). Based on the RNase Y preference for a G at the cleavage site (Figure 11D), it is likely that these fragment 5' ends were also generated by RNase Y. This led to the identification of 127 possible additional RNase Y cleavage sites which were not and could not have been mapped by the RNA end abundance comparison between the WT and Δ*rny* strains.

In addition, we also observed that the interplay of RNase Y and PNPase was crucial for the degradation of premature terminated transcripts derived from riboswitches and T-box elements. The concerted activity of RNase Y and PNPase on regulatory elements was previously observed in both *B. subtilis*[119] and *S. aureus*[160]. The degradation of these premature terminated transcripts could play an important role in recycling of the metabolite bound to the 5' regulatory element, as shown for the *trp* leader (controlling the expression of the tryptophan operon) in *B. subtilis*. In the presence of tryptophan the *trp* RNA binding attenuation protein (TRAP) interacts

with the *trp* leader, promoting transcription termination. The PNPase-mediated degradation of the TRAP-bound leader RNA assures the release of TRAP and therefore maintenance of the *trp* operon repression[385].

In addition, we demonstrated that RNase Y and PNPase activity allows the differential RNA degradation of mRNAs comprised within the same polycistronic transcript (Figure 27). Indeed, we observed that RNase Y processed the *rsmC-cdd-bmpA* transcript operon and PNPase degrades the *rsmC-cdd* isoform, which is less stable than the *bmpA* mRNA (Figure 27). The effect of the differential RNA degradation of this operon at the protein level should be the subject of further investigation. Bacterial operons consist of clustered genes that are co-transcribed from a common promoter, resulting in the production of polycistronic transcripts. The uncoupling of the translation process of the co-transcribed ORFs is fundamental when the protein encoded by one gene is needed at a different level than the protein encoded by the other co-transcribed genes. Processing by an endoRNase within the polycistronic transcript can lead to the differential degradation of the generated RNA products (encompassing a single ORF), allowing to fine-tune the protein levels. The differential expression of genes encoded in the same operon was extensively study for the glycolytic *gapA* operon in *B. subtilis*[123,151,386]. The first gene of operon is *cggR*, encoding the transcriptional repressor of the operon CggR, which is followed by *gapA*, encoding the glyceraldehyde 3-phosphate dehydrogenase. In presence of glucose, the amount of GapA glycolytic enzyme required by the bacterial cell is much higher compared to CggR, which is instead not needed[151]. The differential protein synthesis is obtained by the processing of RNase Y at the end of the *cggR* ORF and subsequent degradation of the RNA product containing the *cggR* ORF by PNPase and stabilization of the RNA product comprising the *gapA* ORF[123,136,387].

4.3.2 RNase Y and YhaM interplay

S. pyogenes YhaM displays a peculiar activity as it trims not specifically most of the RNA 3' ends that derive from transcription termination or endoRNase cleavage. As expected YhaM also trimmed the RNA 3' ends generated by RNase Y processing (Figure 20). In the present thesis, we gained more insight into the activity of this peculiar 3'-to-5' exoRNase in *S. pyogenes*.

First, from the comparison of the RNase Y, PNPase and YhaM targetomes, we observed that YhaM also trims the RNA 3' ends resulting from the trimming of another 3'-to-5' exoRNase (*i.e.* PNPase), after RNase Y processing (Figure 23).

Second, we observed that YhaM trimmed short RNA fragments, produced by RNase Y and that in the absence of YhaM these RNA fragments were not detectable (Figure 25). Based on this observation, we hypothesized that YhaM trimming exerts a protective role, by preventing RNA

degradation performed by PNPase or RNase R. Indeed, both these two enzymes require a ssRNA region at the RNA 3' end to bind and degrade the targets[130,136,374]. A similar function was already proposed in *B. subtilis*, where the YhaM trimming of ssRNA tails at the RNA 3' ends prevents RNase R activity[136].

Some of these short RNA fragments, generated by RNase Y and detectable in WT, were trimmed by both YhaM and PNPase or uniquely by one of the two 3'-to-5' exoRNases. Further characterization of these fragments revealed the presence of putative RNA structure at the 3' fragment region (Figure 24C). While the stable structure could protect the fragments from degradation, we do not exclude the possibility that these short RNA fragments, derived from mRNA degradation, could exert a functional role, as described in *E. coli*. In this bacterium, indeed, stabilized RNA fragments, generated by RNase E processing of mRNAs, act as a source of regulatory sRNAs[32]. These fragments were also described to be structured and bioinformatic prediction revealed that they likely interact with Hfq and/or ProQ[32].

4.4 The role of RNase Y in RNA degradation

RNase Y is often considered to functionally replace RNase E in the bacteria that do not code for this endoRNase[84]. RNase E is an essential enzyme that plays a central role in the initiation of RNA degradation in Gram-negative negative bacteria. Here, we have demonstrated that RNase Y, by cleaving an RNA molecule internally is able to initiate RNA degradation in *S. pyogenes* (Figures 19, 20 and 26). However, based on the detected RNA ends, we concluded that RNase Y is probably not the major endoRNase in this bacterium. We have identified only 320 RNase Y processing positions, by mapping both the RNA 5' and 3' ends (Figures 11 and 18), and additional 127 putative RNase Y processing sites that were detected in the Δ*pnpA* strain. In contrast, mapping of the RNA 5' ends in *S. enterica*, revealed nearly 22000 cleavage sites of RNase E, corresponding to more than one processing event per transcript[30]. It is possible that this high number of RNase E processing sites identified is due to the absence of RNase J1 in this bacterium, which degrades the RNAs in 5'-to-3' direction, allowing therefore to capture numerous RNA 5' ends. The sole RNase in *E. coli*, that was shown *in vitro* to degrade short oligoRNAs in the 5'-to-3' direction is RNase AM[186]. However, this activity has not been characterized *in vivo* yet and the role of this RNase in RNA degradation in *E. coli* is not known.

The number of RNase Y cleavage sites identified in this study is coherent with previous transcriptomic studies in *B. subtilis* and *S. aureus*, in which approximately 100 RNase Y cleavage sites were identified by mapping only the RNA 5' ends[153,160]. The relatively small RNase Y targetome in *S. pyogenes* is consistent with the limited delay in growth measured in the Δ*rny* strain compared to

the WT strain (Figure 7A). This is opposite to *B. subtilis*, where *rny* deletion results in severe phenotypes and changes in the expression of 13% to 23% of the transcriptome[153,162].

Other RNases are likely involved in the initiation of RNA degradation in *S. pyogenes*, including the complex formed by RNases J1 and J2, both essential in this bacterium. Whereas RNase J1 plays an important role in RNA degradation in *B. subtilis*[164], RNase J2 is considered to mainly function as a structural protein in the RNase J1/J2 complex. *S. mutans* is the sole bacterium where RNase J2 was demonstrated to act as an endoRNase *in vivo*[194]. In *S. pyogenes*, RNase J2 was shown to affect transcript stability[195,343] and it is possible that this RNase functions as an endoRNase initiating RNA degradation, as described for *S. mutants*.

A previous study, in M49 *S. pyogenes* strain NZ131, investigated the impact of RNase Y on transcript stability, during growth in a medium poor in carbohydrates and revealed that the 98% of the transcripts were stabilized in the Δ*rny* strain[106]. In this thesis, the RNase Y targetome was determined in M1 *S. pyogenes* strain SF370 grown in a medium rich in carbohydrates at mid-logarithmic growth phase (Figure 7A). Identification of the RNase Y cleavage positions in different conditions (*e.g.* the ones used in the transcriptome stability study) will likely shed light on novel RNase Y direct targets and post-transcriptional mechanisms of gene expression. Interestingly, RNase Y was shown to affect the expression of some genes based on carbon or peptide availability[168]. For instance, the RNase Y-mediated regulation of *speB*, in carbon-poor medium and stationary phase of growth, seems to be influenced by the peptide levels[168]. To date, it is unclear how RNase Y could sense the cells nutritional status and degrade RNAs accordingly.

4.5 General conclusions

To conclude, this thesis provides insights into the activity and functions of RNase Y in *S. pyogenes* (Figure 31). First, we have investigated the RNase Y involvement in the regulation of *speB*, coding for a major virulence factor. With this regard, we have demonstrated that RNase Y, by acting indirectly at transcriptional level, modulates *speB* expression (Figure 31A). RNase Y also processes the *speB* mRNA 5' UTR downstream of a G, giving rise to a putative regulatory sRNA (Figure 31B). Overall, we have established that the major mechanism of *speB* regulation by RNase Y is indirect at transcriptional level. Nonetheless, we do not exclude that the RNase Y processing of the *speB* mRNA plays a more relevant role in *speB* regulation, under specific circumstances (*e.g.* infection, stress etc.) differing from the standard growth conditions used in this study.

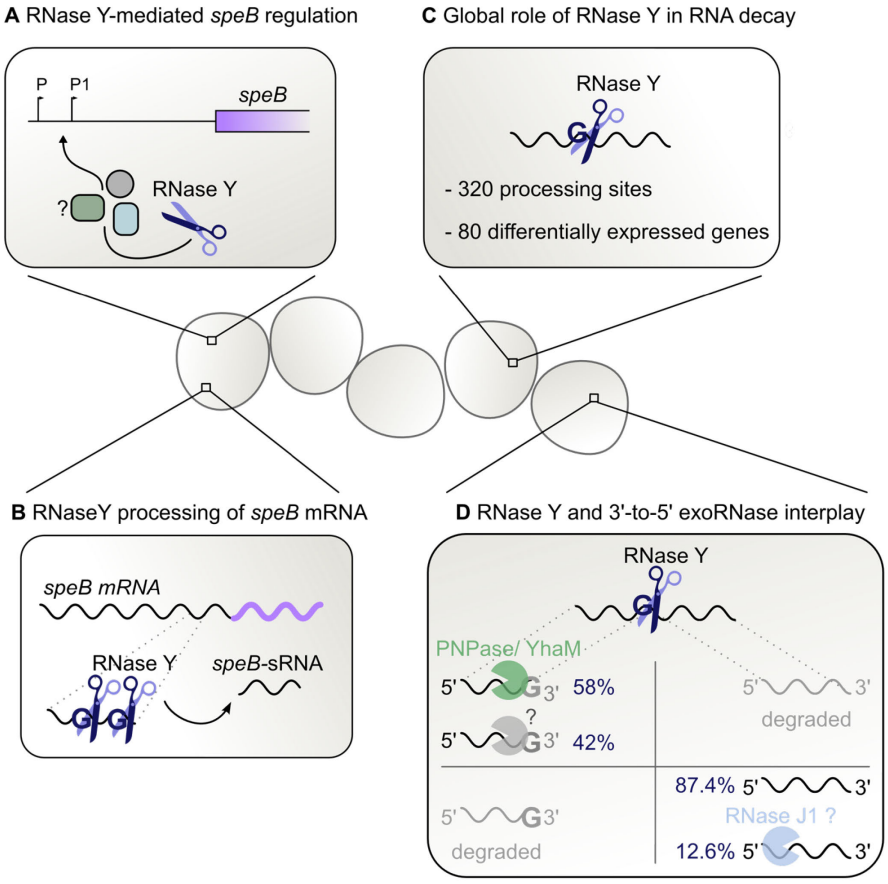


Figure 31. **Study of RNase Y activity in *S. pyogenes*.**
Schematic representation of the main findings of the present thesis. **A.** RNase Y, through unknow intermediary factors, reduces *speB* transcription driven by the P and P1 promoters. **B.** RNase Y processes downstream of a G the *speB* mRNA 5' UTR, leading to the production of a putative sRNA (*speB*-sRNA). **C.** Mapping of the 5' and 3' RNA ends generated by RNase Y uncovered 320 RNase Y processing sites and we observed that RNase Y preferentially cleaves downstream of a G. RNA sequencing revealed that RNase Y affects the expression of 80 genes in *S. pyogenes*. **D.** The detected RNase Y-generated 3' ends were in 52% of the cases trimmed by PNPase or YhaM (green "pacman" symbol) and the remaining 48% of 3' RNA ends were either trimmed by multiple 3'-to-5' exoRNases or by unidentified exoRNases (grey "pacman" symbol). The detected RNase Y-generated RNA 5' ends were in 87.4% of the cases not further trimmed, conversely the remaining 12.6% of detected RNA 5' ends were likely targeted by a 5'-to-3' exoRNase (*i.e.* RNase J1; light blue "pacman" symbol).

Next, we have determined the targetome of RNase Y and gained insight into the substrate specificity of this endoRNase (Figure 31C). RNase Y preferentially cleaves the targets downstream of a G and we have showed that this residue is crucial for the RNase Y processing of the *speB* mRNA 5' UTR.

Lastly, we observed that at least one or both the RNA products deriving from RNase Y processing are degraded by exoRNases, indicating that RNase Y initiates RNA degradation. We employed an RNA sequencing based comparative approach to investigate the interplay of RNase Y with the three 3'-to-5' exoRNases from *S. pyogenes* (Figure 31D). Our data support a model in which RNase Y mainly acts in concert with PNPase in RNA decay.

Overall, we have addressed the complexity of RNase Y role and target specificity in the human pathogen *S. pyogenes*. This study not only gave insight into the machinery and mechanisms of RNA degradation in this human pathogen, but it could pave the way for further understanding of the diversity and function of RNases in different microorganisms.

Bibliography

1. Jacob, F. & Monod, J. Genetic regulatory mechanisms in the synthesis of proteins. *J. Mol. Biol.* **3**, 318–356 (1961).

2. Gilbert, W. & Müller-Hill, B. Isolation of the Lac repressor. *Proc. Natl. Acad. Sci. U. S. A.* **56**, 1891–1898 (1966).

3. Ptashne, M. Isolation of the λ phage repressor. *Proc. Natl. Acad. Sci. U. S. A.* **57**, 306–313 (1967).

4. Tomizawa, J., Itoh, T., Selzer, G. & Som, T. Inhibition of ColE1 RNA primer formation by a plasmid-specified small RNA. *Proc. Natl. Acad. Sci. U. S. A.* **78**, 1421–1425 (1981).

5. Masukata, H. & Tomizawa, J. Control of primer formation for ColE1 plasmid replication: conformational change of the primer transcript. *Cell* **44**, 125–136 (1986).

6. Wagner, E. G. H. & Romby, P. Small RNAs in bacteria and archaea: who they are, what they do, and how they do it. *Adv. Genet.* **90**, 133–208 (2015).

7. Waters, L. S. & Storz, G. Regulatory RNAs in Bacteria. *Cell* **136**, 615–628 (2009).

8. Wagner, E. G. H. & Simons, R. W. Antisense RNA control in bacteria, phages, and plasmids. *Annu. Rev. Microbiol.* **48**, 713–742 (1994).

9. Brantl, S. Antisense-RNA mediated control of plasmid replication - pIP501 revisited. *Plasmid* **78**, 4–16 (2015).

10. Brantl, S. & Jahn, N. sRNAs in bacterial type I and type III toxin-antitoxin systems. *FEMS Microbiol. Rev.* **39**, 413–427 (2015).

11. Masachis, S. & Darfeuille, F. Type I Toxin-Antitoxin systems: regulating toxin expression via Shine-Dalgarno sequence sequestration and small RNA Binding. *Microbiol. Spectr.* **6**, (2018).

12. Gerdes, K., Rasmussen, P. B. & Molin, S. Unique type of plasmid maintenance function: postsegregational killing of plasmid-free cells. *Proc. Natl. Acad. Sci.* **83**, 3116–3120 (1986).

13. Thisted, T. & Gerdes, K. Mechanism of post-segregational killing by the *hok*/*sok* system of plasmid R1. Sok antisense RNA regulates *hok* gene expression indirectly through the overlapping *mok* gene. *J. Mol. Biol.* **223**, 41–54 (1992).

14. Gerdes, K. & Wagner, E. G. H. RNA antitoxins. *Curr. Opin. Microbiol.* **10**, 117–124 (2007).

15. Selinger, D. W., Saxena, R. M., Cheung, K. J., Church, G. M. & Rosenow, C. Global RNA half-life analysis in *Escherichia coli* reveals positional patterns of transcript degradation. *Genome Res.* **13**, 216–223 (2003).

16. Lasa, I. *et al.* Genome-wide antisense transcription drives mRNA processing in bacteria. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 20172–20177 (2011).

17. Thomason, M. K. *et al.* Global transcriptional start site mapping using differential RNA sequencing reveals novel antisense RNAs in *Escherichia coli*. *J. Bacteriol.* **197**, 18–28 (2015).

18. Wade, J. T. & Grainger, D. C. Pervasive transcription: illuminating the dark matter of bacterial transcriptomes. *Nat. Rev. Microbiol.* **12**, 647–653 (2014).

19. Georg, J. & Hess, W. R. Widespread antisense transcription in prokaryotes. *Microbiol. Spectr.* **6**,(4) (2018).

20. Toledo-Arana, A. *et al.* The *Listeria* transcriptional landscape from saprophytism to virulence. *Nature* **459**, 950–956 (2009).

21. Sesto, N., Wurtzel, O., Archambaud, C., Sorek, R. & Cossart, P. The excludon: a new concept in bacterial antisense RNA-mediated gene regulation. *Nat. Rev. Microbiol.* **11**, 75–82 (2013).

22. Opdyke, J. A., Kang, J.-G. & Storz, G. GadY, a small-RNA regulator of acid response genes in *Escherichia coli*. *J. Bacteriol.* **186**, 6698–6705 (2004).

23. Opdyke, J. A., Fozo, E. M., Hemm, M. R. & Storz, G. RNase III participates in GadY-dependent cleavage of the *gadX-gadW* mRNA. *J. Mol. Biol.* **406**, 29–43 (2011).

24. Tramonti, A., De Canio, M. & De Biase, D. GadX/GadW-dependent regulation of the *Escherichia coli* acid fitness island: transcriptional control at the *gadY-gadW* divergent promoters and identification of four novel 42 bp GadX/GadW-specific binding sites. *Mol. Microbiol.* **70**, 965–982 (2008).

25. Sáenz-Lahoya, S. *et al.* Noncontiguous operon is a genetic organization for coordinating bacterial gene expression. *Proc. Natl. Acad. Sci.* **116**, 1733–1738 (2019).

26. Argaman, L. *et al.* Novel small RNA-encoding genes in the intergenic regions of *Escherichia coli*. *Curr. Biol. CB* **11**, 941–950 (2001).

27. Wassarman, K. M., Repoila, F., Rosenow, C., Storz, G. & Gottesman, S. Identification of novel small RNAs using comparative genomics and microarrays. *Genes Dev.* **15**, 1637–1651 (2001).

28. Loh, E. *et al.* A *trans*-acting riboswitch controls expression of the virulence regulator PrfA in *Listeria monocytogenes*. *Cell* **139**, 770–779 (2009).

29. Chao, Y. & Vogel, J. A 3' UTR-derived small RNA provides the regulatory noncoding Arm of the inner membrane stress response. *Mol. Cell* **61**, 352–363 (2016).

30. Chao, Y. *et al. In vivo* cleavage map illuminates the central role of RNase E in coding and non-coding RNA pathways. *Mol. Cell* **65**, 39–51 (2017).

31. De Lay, N. R. & Garsin, D. A. The unmasking of "junk" RNA reveals novel sRNAs: from processed RNA fragments to marooned riboswitches. *Curr. Opin. Microbiol.* **30**, 16–21 (2016).

32. Dar, D. & Sorek, R. Bacterial noncoding RNAs excised from within protein-coding transcripts. *mBio* **9**, 01730–18 (2018).

33. Carrier, M.-C., Lalaouna, D. & Massé, E. Broadening the definition of bacterial small RNAs: characteristics and mechanisms of action. *Annu. Rev. Microbiol.* **72**, 141–161 (2018).

34. Papenfort, K. & Vanderpool, C. K. Target activation by regulatory RNAs in bacteria. *FEMS Microbiol. Rev.* **39**, 362–378 (2015).

35. Pfeiffer, V., Papenfort, K., Lucchini, S., Hinton, J. C. D. & Vogel, J. Coding sequence targeting by MicC RNA reveals bacterial mRNA silencing downstream of translational initiation. *Nat. Struct. Mol. Biol.* **16**, 840–846 (2009).

36. Bandyra, K. J. *et al.* The seed region of a small RNA drives the controlled destruction of the target mRNA by the endoribonuclease RNase E. *Mol. Cell* **47**, 943–953 (2012).

37. Durand, S., Braun, F., Helfer, A.-C., Romby, P. & Condon, C. sRNA-mediated activation of gene expression by inhibition of 5'-3' exonucleolytic mRNA degradation. *eLife* **6**, (2017).

38. Miyakoshi, M., Chao, Y. & Vogel, J. Cross talk between ABC transporter mRNAs via a target mRNA-derived sponge of the GcvB small RNA. *EMBO J.* **34**, 1478–1492 (2015).

39. Lalaouna, D. *et al.* A 3' external transcribed spacer in a tRNA transcript acts as a sponge for small RNAs to prevent transcriptional noise. *Mol. Cell* **58**, 393–405 (2015).

40. Breaker, R. R. Riboswitches and translation control. *Cold Spring Harb. Perspect. Biol.* **10**, (2018).

41. Ignatov, D. & Johansson, J. RNA-mediated signal perception in pathogenic bacteria. *Wiley Interdiscip. Rev. RNA* **8**, e1429 (2017).

42. Serganov, A. & Nudler, E. A decade of riboswitches. *Cell* **152**, 17–24 (2013).

43. Green, N. J., Grundy, F. J. & Henkin, T. M. The T box mechanism: tRNA as a regulatory molecule. *FEBS Lett.* **584**, 318–324 (2010).

44. Gutiérrez-Preciado, A., Henkin, T. M., Grundy, F. J., Yanofsky, C. & Merino, E. Biochemical features and functional implications of the RNA-based T-box regulatory mechanism. *Microbiol. Mol. Biol. Rev.* **73**, 36–61 (2009).

45. Romeo, T., Vakulskas, C. A. & Babitzke, P. Posttranscriptional regulation on a global scale: form and function of Csr/Rsm systems. *Environ. Microbiol.* **15**, 313–324 (2013).

46. Lee, S. Y., Bailey, S. C. & Apirion, D. Small stable RNAs from *Escherichia coli*: evidence for the existence of new molecules and for a new ribonucleoprotein particle containing 6S RNA. *J. Bacteriol.* **133**, 1015–1023 (1978).

47. Wassarman, K. M. & Storz, G. 6S RNA regulates *E. coli* RNA polymerase activity. *Cell* **101**, 613–623 (2000).

48. Wassarman, K. M. 6S RNA, a global regulator of transcription. *Microbiol. Spectr.* **6**, (2018).

49. Gruber, T. M. & Gross, C. A. Multiple sigma subunits and the partitioning of bacterial transcription space. *Annu. Rev. Microbiol.* **57**, 441–466 (2003).

50. Cavanagh, A. T. & Wassarman, K. M. 6S RNA, a global regulator of transcription in *Escherichia coli*, *Bacillus subtilis*, and beyond. *Annu. Rev. Microbiol.* **68**, 45–60 (2014).

51. Blower, T. R. *et al.* Identification and classification of bacterial Type III toxin–antitoxin systems encoded in chromosomal and plasmid genomes. *Nucleic Acids Res.* **40**, 6158–6173 (2012).

52. Hille, F. *et al.* The Biology of CRISPR-Cas: backward and forward. *Cell* **172**, 1239–1259 (2018).

53. Deltcheva, E. *et al.* CRISPR RNA maturation by *trans*-encoded small RNA and host factor RNase III. *Nature* **471**, 602–607 (2011).

54. Le Rhun, A., Escalera-Maurer, A., Bratovič, M. & Charpentier, E. CRISPR-Cas in *Streptococcus pyogenes*. *RNA Biol.* **16**, 380–389 (2019).

55. Sampson, T. R., Saroj, S. D., Llewellyn, A. C., Tzeng, Y.-L. & Weiss, D. S. A CRISPR/Cas system mediates bacterial innate immune evasion and virulence. *Nature* **497**, 254–257 (2013).

56. Louwen, R., Staals, R. H. J., Endtz, H. P., van Baarlen, P. & van der Oost, J. The role of CRISPR-Cas systems in virulence of pathogenic bacteria. *Microbiol. Mol. Biol. Rev. MMBR* **78**, 74–88 (2014).

57. Ma, K. *et al. cas9* enhances bacterial virulence by repressing the *reg*R transcriptional regulator in *Streptococcus agalactiae*. *Infect. Immun.* **86**, (2018).

58. Dugar, G. *et al.* CRISPR RNA-dependent binding and cleavage of endogenous RNAs by the *Campylobacter jejuni* Cas9. *Mol. Cell* **69**, 893-905.e7 (2018).

59. Ratner, H. K. *et al.* Catalytically active Cas9 mediates transcriptional interference to facilitate bacterial virulence. *Mol. Cell* **75**, 498-510.e5 (2019).

60. Holmqvist, E. & Vogel, J. RNA-binding proteins in bacteria. *Nat. Rev. Microbiol.* **16**, 601 (2018).

61. Updegrove, T. B., Zhang, A. & Storz, G. Hfq: the flexible RNA matchmaker. *Curr. Opin. Microbiol.* **30**, 133–138 (2016).

62. Sauer, E. & Weichenrieder, O. Structural basis for RNA 3'-end recognition by Hfq. *Proc. Natl. Acad. Sci.* **108**, 13065–13070 (2011).

63. Mikulecky, P. J. *et al.* *Escherichia coli* Hfq has distinct interaction surfaces for DsrA, *rpoS* and poly(A) RNAs. *Nat. Struct. Mol. Biol.* **11**, 1206–1214 (2004).

64. Link, T. M., Valentin-Hansen, P. & Brennan, R. G. Structure of *Escherichia coli* Hfq bound to polyriboadenylate RNA. *Proc. Natl. Acad. Sci.* **106**, 19292–19297 (2009).

65. Morita, T., Maki, K. & Aiba, H. RNase E-based ribonucleoprotein complexes: mechanical basis of mRNA destabilization mediated by bacterial noncoding RNAs. *Genes Dev.* **19**, 2176–2186 (2005).

66. Bruce, H. A. *et al.* Analysis of the natively unstructured RNA/protein-recognition core in the *Escherichia coli* RNA degradosome and its interactions with regulatory RNA/Hfq complexes. *Nucleic Acids Res.* **46**, 387–402 (2018).

67. Romilly, C. *et al.* Current knowledge on regulatory RNAs and their machineries in *Staphylococcus aureus*. *RNA Biol.* **9**, 402–413 (2012).

68. Le Rhun, A. & Charpentier, E. Small RNAs in streptococci. *RNA Biol.* **9**, 414–426 (2012).

69. Dambach, M., Irnov, I. & Winkler, W. C. Association of RNAs with *Bacillus subtilis* Hfq. *PLOS ONE* **8**, e55156 (2013).

70. Nielsen, J. S. *et al.* Defining a role for Hfq in Gram-positive bacteria: evidence for Hfq-dependent antisense regulation in *Listeria monocytogenes*. *Nucleic Acids Res.* **38**, 907–919 (2010).

71. Smaldone, G. T., Antelmann, H., Gaballa, A. & Helmann, J. D. The FsrA sRNA and FbpB protein mediate the iron-dependent induction of the *Bacillus subtilis* LutABC iron-sulfur-containing oxidases. *J. Bacteriol.* **194**, 2586–2593 (2012).

72. Attaiech, L., Glover, J. N. M. & Charpentier, X. RNA chaperones step out of Hfq's shadow. *Trends Microbiol.* **25**, 247–249 (2017).

73. Smirnov, A. *et al.* Grad-seq guides the discovery of ProQ as a major small RNA-binding protein. *Proc. Natl. Acad. Sci.* **113**, 11591–11596 (2016).

74. Olejniczak, M. & Storz, G. ProQ/FinO-domain proteins: another ubiquitous family of RNA matchmakers? *Mol. Microbiol.* **104**, 905–915 (2017).

75. Attaiech, L. *et al.* Silencing of natural transformation by an RNA chaperone and a multitarget small RNA. *Proc. Natl. Acad. Sci.* **113**, 8813–8818 (2016).

76. Mark Glover, J. N. *et al.* The FinO family of bacterial RNA chaperones. *Plasmid* **78**, 79–87 (2015).

77. Khemici, V. & Linder, P. RNA helicases in bacteria. *Curr. Opin. Microbiol.* **30**, 58–66 (2016).

78. Kaberdin, V. R. & Bläsi, U. Bacterial helicases in post-transcriptional control. *Biochim. Biophys. Acta BBA - Gene Regul. Mech.* **1829**, 878–883 (2013).

79. Mitra, P., Ghosh, G., Hafeezunnisa, M. & Sen, R. Rho protein: roles and mechanisms. *Annu. Rev. Microbiol.* **71**, 687–709 (2017).

80. Wang, X. *et al.* Processing generates 3' ends of RNA masking transcription termination events in prokaryotes. *Proc. Natl. Acad. Sci.* **116**, 4440–4445 (2019).

81. Py, B., Higgins, C. F., Krisch, H. M. & Carpousis, A. J. A DEAD-box RNA helicase in the *Escherichia coli* RNA degradosome. *Nature* **381**, 169–172 (1996).

82. Khemici, V., Toesca, I., Poljak, L., Vanzo, N. F. & Carpousis, A. J. The RNase E of *Escherichia coli* has at least two binding sites for DEAD-box RNA helicases: functional replacement of RhlB by RhlE. *Mol. Microbiol.* **54**, 1422–1430 (2004).

83. Vakulskas, C. A. *et al.* Global effects of the DEAD-box RNA helicase DeaD (CsdA) on gene expression over a broad range of temperatures. *Mol. Microbiol.* **92**, 945–958 (2014).

84. Durand, S. & Condon, C. RNases and helicases in Gram-positive bacteria. *Microbiol. Spectr.* **6(2)**, (2018).

85. Pandiani, F. *et al.* Differential involvement of the five RNA helicases in adaptation of *Bacillus cereus* ATCC 14579 to low growth temperatures. *Appl. Environ. Microbiol.* **76**, 6692–6697 (2010).

86. Netterling, S., Bäreclev, C., Vaitkevicius, K. & Johansson, J. RNA helicase important for *Listeria monocytogenes* hemolytic activity and virulence factor expression. *Infect. Immun.* **84**, 67–76 (2016).

87. Lehnik-Habrink, M. *et al.* DEAD-Box RNA helicases in *Bacillus subtilis* have multiple functions and act independently from each other. *J. Bacteriol.* **195**, 534–544 (2013).

88. Lehnik-Habrink, M. *et al.* The RNA degradosome in *Bacillus subtilis*: identification of CshA as the major RNA helicase in the multiprotein complex. *Mol. Microbiol.* **77**, 958–971 (2010).

89. Oun, S. *et al.* The CshA DEAD-box RNA helicase is important for quorum sensing control in *Staphylococcus aureus*. *RNA Biol.* **10**, 157–165 (2013).

90. Bronesky, D. *et al. Staphylococcus aureus* RNAIII and its regulon link quorum sensing, stress responses, metabolic adaptation, and regulation of virulence gene expression. *Annu. Rev. Microbiol.* **70**, 299–316 (2016).

91. Anderson, K. L. & Dunman, P. M. Messenger RNA turnover processes in *Escherichia coli*, *Bacillus subtilis*, and emerging studies in *Staphylococcus aureus*. *Int. J. Microbiol.* **2009**, (2009).

92. Durand, S., Tomasini, A., Braun, F., Condon, C. & Romby, P. sRNA and mRNA turnover in Gram-positive bacteria. *FEMS Microbiol. Rev.* **39**, 316–330 (2015).

93. Celesnik, H., Deana, A. & Belasco, J. G. Initiation of RNA decay in *Escherichia coli* by 5' pyrophosphate removal. *Mol. Cell* **27**, 79–90 (2007).

94. Richards, J. *et al.* An RNA pyrophosphohydrolase triggers 5'-exonucleolytic degradation of mRNA in *Bacillus subtilis*. *Mol. Cell* **43**, 940–949 (2011).

95. Cahová, H., Winz, M.-L., Höfer, K., Nübel, G. & Jäschke, A. NAD captureSeq indicates NAD as a bacterial cap for a subset of regulatory RNAs. *Nature* **519**, 374–377 (2015).

96. Frindert, J. *et al.* Identification, biosynthesis, and decapping of NAD-capped RNAs in *B. subtilis*. *Cell Rep.* **24**, 1890-1901.e8 (2018).

97. Sharp, J. S. & Bechhofer, D. H. Effect of 5'-proximal elements on decay of a model mRNA in *Bacillus subtilis*. *Mol. Microbiol.* **57**, 484–495 (2005).

98. Hambraeus, G., von Wachenfeldt, C. & Hederstedt, L. Genome-wide survey of mRNA half-lives in *Bacillus subtilis* identifies extremely stable mRNAs. *Mol. Genet. Genomics MGG* **269**, 706–714 (2003).

99. McLaren, R. S., Newbury, S. F., Dance, G. S., Causton, H. C. & Higgins, C. F. mRNA degradation by processive 3'-5' exoribonucleases in vitro and the implications for prokaryotic mRNA decay *in vivo*. *J. Mol. Biol.* **221**, 81–95 (1991).

100. Agaisse, H. & Lereclus, D. STAB-SD: a Shine-Dalgarno sequence in the 5' untranslated region is a determinant of mRNA stability. *Mol. Microbiol.* **20**, 633–643 (1996).

101. Braun, F., Le Derout, J. & Régnier, P. Ribosomes inhibit an RNase E cleavage which induces the decay of the *rpsO* mRNA of *Escherichia coli*. *EMBO J.* **17**, 4790–4797 (1998).

102. Braun, F., Durand, S. & Condon, C. Initiating ribosomes and a 5'/3'-UTR interaction control ribonuclease action to tightly couple *B. subtilis hbs* mRNA stability with translation. *Nucleic Acids Res.* **45**, 11386–11400 (2017).

103. Moffitt, J. R., Pandey, S., Boettiger, A. N., Wang, S. & Zhuang, X. Spatial organization shapes the turnover of a bacterial transcriptome. *eLife* **5**, e13065 (2016).

104. Deutscher, M. P. & Reuven, N. B. Enzymatic basis for hydrolytic versus phosphorolytic mRNA degradation in *Escherichia coli* and *Bacillus subtilis*. *Proc. Natl. Acad. Sci. U. S. A.* **88**, 3277–3280 (1991).

105. Roberts, C. *et al.* Characterizing the effect of the *Staphylococcus aureus* virulence factor Regulator, SarA, on log-phase mRNA half-lives. *J. Bacteriol.* **188**, 2593–2603 (2006).

106. Chen, Z., Itzek, A., Malke, H., Ferretti, J. J. & Kreth, J. Multiple roles of RNase Y in *Streptococcus pyogenes* mRNA processing and degradation. *J. Bacteriol.* **195**, 2585–2594 (2013).

107. Deutscher, M. P. Degradation of stable RNA in bacteria. *J. Biol. Chem.* **278**, 45041–45044 (2003).

108. Laalami, S., Zig, L. & Putzer, H. Initiation of mRNA decay in bacteria. *Cell. Mol. Life Sci. CMLS* **71**, 1799–1828 (2014).

109. Bechhofer, D. H. & Deutscher, M. P. Bacterial ribonucleases and their roles in RNA metabolism. *Crit. Rev. Biochem. Mol. Biol.* **54**, 242–300 (2019).

110. Deana, A., Celesnik, H. & Belasco, J. G. The bacterial enzyme RppH triggers messenger RNA degradation by 5' pyrophosphate removal. *Nature* **451**, 355–358 (2008).

111. Luciano, D. J., Vasilyev, N., Richards, J., Serganov, A. & Belasco, J. G. A novel RNA phosphorylation state enables 5' end-dependent degradation in *Escherichia coli*. *Mol. Cell* **67**, 44-54.e6 (2017).

112. Luciano, D. J., Vasilyev, N., Richards, J., Serganov, A. & Belasco, J. G. Importance of a diphosphorylated intermediate for RppH-dependent RNA degradation. *RNA Biol.* **15**, 703–706 (2018).

113. Mackie, G. A. Ribonuclease E is a 5'-end-dependent endonuclease. *Nature* **395**, 720–723 (1998).

114. Koslover, D. J. *et al.* The crystal structure of the *Escherichia coli* RNase E apoprotein and a mechanism for RNA degradation. *Structure* **16**, 1238–1244 (2008).

115. Richards, J. & Belasco, J. G. Obstacles to scanning by RNase E govern bacterial mRNA lifetimes by hindering access to distal cleavage sites. *Mol. Cell* **74**, 284-295.e5 (2019).

116. Condon, C. & Putzer, H. The phylogenetic distribution of bacterial ribonucleases. *Nucleic Acids Res.* **30**, 5339–5346 (2002).

117. Mathy, N. *et al.* 5'-to-3' exoribonuclease activity in bacteria: role of RNase J1 in rRNA maturation and 5' stability of mRNA. *Cell* **129**, 681–692 (2007).

118. Yao, S., Sharp, J. S. & Bechhofer, D. H. *Bacillus subtilis* RNase J1 endonuclease and 5' exonuclease activities in the turnover of DeltaermC mRNA. *RNA* **15**, 2331–2339 (2009).

119. Shahbabian, K., Jamalli, A., Zig, L. & Putzer, H. RNase Y, a novel endoribonuclease, initiates riboswitch turnover in *Bacillus subtilis*. *EMBO J.* **28**, 3523–3533 (2009).

120. Hui, M. P., Foley, P. L. & Belasco, J. G. Messenger RNA degradation in bacterial cells. *Annu. Rev. Genet.* **48**, 537–559 (2014).

121. Clarke, J. E., Kime, L., Romero A, D. & McDowall, K. J. Direct entry by RNase E is a major pathway for the degradation and processing of RNA in *Escherichia coli*. *Nucleic Acids Res.* **42**, 11733–11751 (2014).

122. Yao, S. & Bechhofer, D. H. Initiation of decay of *Bacillus subtilis rpsO* mRNA by endoribonuclease RNase Y. *J. Bacteriol.* **192**, 3279–3286 (2010).

123. Liu, B. *et al.* Global analysis of mRNA decay intermediates in *Bacillus subtilis* wild-type and polynucleotide phosphorylase-deletion strains. *Mol. Microbiol.* **94**, 41–55 (2014).

124. Kinscherf, T. G. & Apirion, D. Polynucleotide phosphorylase can participate in decay of mRNA in *Escherichia coli* in the absence of ribonuclease II. *Mol. Gen. Genet. MGG* **139**, 357–362 (1975).

125. Donovan, W. P. & Kushner, S. R. Polynucleotide phosphorylase and ribonuclease II are required for cell viability and mRNA turnover in *Escherichia coli* K-12. *Proc. Natl. Acad. Sci.* **83**, 120–124 (1986).

126. Mohanty, B. K. & Kushner, S. R. Genomic analysis in *Escherichia coli* demonstrates differential roles for polynucleotide phosphorylase and RNase II in mRNA abundance and decay. *Mol. Microbiol.* **50**, 645–658 (2003).

127. Liao, H., Liu, M. & Guo, X. The special existences: nanoRNA and nanoRNase. *Microbiol. Res.* **207**, 134–139 (2018).

128. Niyogi, S. K. & Datta, A. K. A novel oligoribonuclease of *Escherichia coli*. I. Isolation and properties. *J. Biol. Chem.* **250**, 7307–7312 (1975).

129. Datta, A. K. & Niyogi, K. A novel oligoribonuclease of *Escherichia coli*. II. Mechanism of action. *J. Biol. Chem.* **250**, 7313–7319 (1975).

130. Cheng, Z.-F. & Deutscher, M. P. An important role for RNase R in mRNA decay. *Mol. Cell* **17**, 313–318 (2005).

131. Chen, X. *et al.* An RNA degradation machine sculpted by Ro autoantigen and noncoding RNA. *Cell* **153**, 166–177 (2013).

132. Liou, G.-G., Chang, H.-Y., Lin, C.-S. & Lin-Chao, S. DEAD box RhlB RNA helicase physically associates with exoribonuclease PNPase to degrade double-stranded RNA independent of the degradosome-assembling region of RNase E. *J. Biol. Chem.* **277**, 41157–41162 (2002).

133. O'Hara, E. B. *et al.* Polyadenylylation helps regulate mRNA decay in *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* **92**, 1807–1811 (1995).

134. Joanny, G. *et al.* Polyadenylation of a functional mRNA controls gene expression in *Escherichia coli*. *Nucleic Acids Res.* **35**, 2494–2502 (2007).

135. Vincent, H. A. & Deutscher, M. P. Substrate recognition and catalysis by the exoribonuclease RNase R. *J. Biol. Chem.* **281**, 29769–29775 (2006).

136. Oussenko, I. A., Abe, T., Ujiie, H., Muto, A. & Bechhofer, D. H. Participation of 3'-to-5' exoribonucleases in the turnover of *Bacillus subtilis* mRNA. *J. Bacteriol.* **187**, 2758–2767 (2005).

137. DiChiara, J. M., Liu, B., Figaro, S., Condon, C. & Bechhofer, D. H. Mapping of internal monophosphate 5' ends of *Bacillus subtilis* messenger RNAs and ribosomal RNAs in wild-type and ribonuclease-mutant strains. *Nucleic Acids Res.* **44**, 3373–3389 (2016).

138. Campos-Guillén, J., Bralley, P., Jones, G. H., Bechhofer, D. H. & Olmedo-Alvarez, G. Addition of poly(A) and heteropolymeric 3' ends in *Bacillus subtilis* wild-type and polynucleotide phosphorylase-deficient strains. *J. Bacteriol.* **187**, 4698–4706 (2005).

139. Fang, M. *et al.* Degradation of nanoRNA is performed by multiple redundant RNases in *Bacillus subtilis*. *Nucleic Acids Res.* **37**, 5114–5125 (2009).

140. Deutscher, M. P. Degradation of RNA in bacteria: comparison of mRNA and stable RNA. *Nucleic Acids Res.* **34**, 659–666 (2006).

141. Deutscher, M. P. Maturation and degradation of ribosomal RNA in bacteria. *Prog. Mol. Biol. Transl. Sci.* **85**, 369–391 (2009).

142. Bram, R. J., Young, R. A. & Steitz, J. A. The ribonuclease III site flanking 23S sequences in the 30S ribosomal precursor RNA of *E. coli*. *Cell* **19**, 393–401 (1980).

143. Young, R. A. & Steitz, J. A. Complementary sequences 1700 nucleotides apart form a ribonuclease III cleavage site in *Escherichia coli* ribosomal precursor RNA. *Proc. Natl. Acad. Sci. U. S. A.* **75**, 3593–3597 (1978).

144. Herskovitz, M. A. & Bechhofer, D. H. Endoribonuclease RNase III is essential in *Bacillus subtilis*. *Mol. Microbiol.* **38**, 1027–1033 (2000).

145. Robertson, H. D., Altman, S. & Smith, J. D. Purification and properties of a specific *Escherichia coli* ribonuclease which cleaves a tyrosine transfer ribonucleic acid precursor. *J. Biol. Chem.* **247**, 5243–5251 (1972).

146. Gardiner, K. & Pace, N. R. RNase P of *Bacillus subtilis* has a RNA component. *J. Biol. Chem.* **255**, 7507–7509 (1980).

147. Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N. & Altman, S. The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell* **35**, 849–857 (1983).

148. Lundberg, U. & Altman, S. Processing of the precursor to the catalytic RNA subunit of RNase P from *Escherichia coli*. *RNA* **1**, 327–334 (1995).

149. Gilet, L., DiChiara, J. M., Figaro, S., Bechhofer, D. H. & Condon, C. Small stable RNA maturation and turnover in *Bacillus subtilis*. *Mol. Microbiol.* **95**, 270–282 (2015).

150. Nilsson, P., Naureckiene, S. & Uhlin, B. E. Mutations affecting mRNA processing and fimbrial biogenesis in the *Escherichia coli pap* operon. *J. Bacteriol.* **178**, 683–690 (1996).

151. Ludwig, H. *et al.* Transcription of glycolytic genes and operons in *Bacillus subtilis*: evidence for the presence of multiple levels of control of the *gapA* operon. *Mol. Microbiol.* **41**, 409–422 (2001).

152. Mäder, U., Hennig, S., Hecker, M. & Homuth, G. Transcriptional organization and posttranscriptional regulation of the *Bacillus subtilis* branched-chain amino acid biosynthesis genes. *J. Bacteriol.* **186**, 2240–2252 (2004).

153. DeLoughery, A., Lalanne, J.-B., Losick, R. & Li, G.-W. Maturation of polycistronic mRNAs by the endoribonuclease RNase Y and its associated Y-complex in *Bacillus subtilis*. *Proc. Natl. Acad. Sci.* **115**, E5585–E5594 (2018).

154. Kim, K. & Lee, Y. Regulation of 6S RNA biogenesis by switching utilization of both sigma factors and endoribonucleases. *Nucleic Acids Res.* **32**, 6057–6068 (2004).

155. Fadouloglou, V. E. *et al.* Maturation of 6S regulatory RNA to a highly elongated structure. *FEBS J.* **282**, 4548–4564 (2015).

156. Durand, S. *et al.* A nitric oxide regulated small RNA controls expression of genes involved in redox homeostasis in *Bacillus subtilis*. *PLoS Genet.* **11**, e1004957 (2015).

157. Lehnik-Habrink, M. *et al.* RNase Y in *Bacillus subtilis*: a natively disordered protein that is the functional equivalent of RNase E from *Escherichia coli*. *J. Bacteriol.* **193**, 5431–5441 (2011).

158. Aït-Bara, S. & Carpousis, A. J. RNA degradosomes in bacteria and chloroplasts: classification, distribution and evolution of RNase E homologs. *Mol. Microbiol.* **97**, 1021–1135 (2015).

159. Commichau, F. M. *et al.* Novel activities of glycolytic enzymes in *Bacillus subtilis*: interactions with essential proteins involved in mRNA processing. *Mol. Cell. Proteomics MCP* **8**, 1350–1360 (2009).

160. Khemici, V., Prados, J., Linder, P. & Redder, P. Decay-initiating endoribonucleolytic cleavage by RNase Y is kept under tight control via sequence preference and sub-cellular localisation. *PLoS Genet* **11**, e1005577 (2015).

161. Chen, Z., Mashburn-Warren, L., Merritt, J., Federle, M. J. & Kreth, J. Interference of a *speB* 5' untranslated region partial deletion with mRNA degradation in *Streptococcus pyogenes*. *Mol. Oral Microbiol.* **32**, 390–403 (2017).

162. Figaro, S. *et al. Bacillus subtilis* mutants with knockouts of the genes encoding ribonucleases RNase Y and RNase J1 are viable, with major defects in cell morphology, sporulation, and competence. *J. Bacteriol.* **195**, 2340–2348 (2013).

163. DeLoughery, A., Dengler, V., Chai, Y. & Losick, R. Biofilm formation by *Bacillus subtilis* requires an endoribonuclease-containing multisubunit complex that controls mRNA levels for the matrix gene repressor SinR. *Mol. Microbiol.* **99**, 425–437 (2016).

164. Durand, S., Gilet, L. & Condon, C. The essential function of *B. subtilis* RNase III is to silence foreign toxin genes. *PLoS Genet.* **8**, e1003181 (2012).

165. Laalami, S. *et al. Bacillus subtilis* RNase Y activity *in vivo* analysed by tiling microarrays. *PloS One* **8**, e54062 (2013).

166. Obana, N., Nakamura, K. & Nomura, N. Role of RNase Y in *Clostridium perfringens* mRNA decay and processing. *J. Bacteriol.* **199**, e00703-16 (2017).

167. Marincola, G. *et al.* RNase Y of *Staphylococcus aureus* and its role in the activation of virulence genes. *Mol. Microbiol.* **85**, 817–832 (2012).

168. Kang, S. O., Caparon, M. G. & Cho, K. H. Virulence gene regulation by CvfA, a putative RNase: the CvfA-Enolase complex in *Streptococcus pyogenes* links nutritional stress, growth-phase control, and virulence gene expression. *Infect. Immun.* **78**, 2754–2767 (2010).

169. Marincola, G. & Wolz, C. Downstream element determines RNase Y cleavage of the *saePQRS* operon in *Staphylococcus aureus*. *Nucleic Acids Res.* **45**, 5980–5994 (2017).

170. Tortosa, P., Albano, M. & Dubnau, D. Characterization of *ylbF*, a new gene involved in competence development and sporulation in *Bacillus subtilis*. *Mol. Microbiol.* **35**, 1110–1119 (2000).

171. Branda, S. S. *et al.* Genes involved in formation of structured multicellular communities by *Bacillus subtilis*. *J. Bacteriol.* **186**, 3970–3979 (2004).

172. Carabetta, V. J. *et al.* A complex of YlbF, YmcA and YaaT regulates sporulation, competence and biofilm formation by accelerating the phosphorylation of Spo0A. *Mol. Microbiol.* **88**, 283–300 (2013).

173. DeLoughery, A., Dengler, V., Chai, Y. & Losick, R. Biofilm formation by *Bacillus subtilis* requires an endoribonuclease-containing multisubunit complex that controls mRNA levels for the matrix gene repressor SinR. *Mol. Microbiol.* **99**, 425–437 (2016).

174. Adusei-Danso, F. *et al.* Structure-function studies of the *Bacillus subtilis* Ric proteins identify the Fe-S cluster-ligating residues and their roles in development and RNA processing. *mBio* **10**, (2019).

175. Bruscella, P., Shahbabian, K., Laalami, S. & Putzer, H. RNase Y is responsible for uncoupling the expression of translation factor IF3 from that of the ribosomal proteins L35 and L20 in *Bacillus subtilis*. *Mol. Microbiol.* **81**, 1526–1541 (2011).

176. Kaito, C. *et al.* Silkworm pathogenic bacteria infection model for identification of novel virulence genes. *Mol. Microbiol.* **56**, 934–944 (2005).

177. Nagata, M., Kaito, C. & Sekimizu, K. Phosphodiesterase activity of CvfA is required for virulence in *Staphylococcus aureus*. *J. Biol. Chem.* **283**, 2176–2184 (2008).

178. Kang, S. O. *et al.* Thermoregulation of capsule production by *Streptococcus pyogenes*. *PLoS ONE* **7**, e37367 (2012).

179. Court, D. L. *et al.* RNase III: genetics and function; structure and mechanism. *Annu. Rev. Genet.* **47**, 405–431 (2013).

180. Nicholson, A. W. Ribonuclease III mechanisms of double-stranded RNA cleavage. *Wiley Interdiscip. Rev. RNA* **5**, 31–48 (2014).

181. Robertson, H. D. *Escherichia coli* ribonuclease III cleavage sites. *Cell* **30**, 669–672 (1982).

182. Calin-Jageman, I. & Nicholson, A. W. RNA structure-dependent uncoupling of substrate recognition and cleavage by *Escherichia coli* ribonuclease III. *Nucleic Acids Res.* **31**, 2381–2392 (2003).

183. Stead, M. B. *et al.* Analysis of *Escherichia coli* RNase E and RNase III activity *in vivo* using tiling microarrays. *Nucleic Acids Res.* **39**, 3188–3203 (2011).

184. Lioliou, E. *et al.* Global regulatory functions of the *Staphylococcus aureus* endoribonuclease III in gene expression. *PLoS Genet* **8**, e1002782 (2012).

185. Le Rhun, A. *et al.* Identification of endoribonuclease specific cleavage positions reveals novel targets of RNase III in *Streptococcus pyogenes*. *Nucleic Acids Res.* **45**, 2329–2340 (2017).

186. Ghodge, S. V. & Raushel, F. M. Discovery of a previously unrecognized ribonuclease from *Escherichia coli* that hydrolyzes 5'-Phosphorylated fragments of RNA. *Biochemistry* **54**, 2911–2918 (2015).

187. Dorléans, A. *et al.* Molecular basis for the recognition and cleavage of RNA by the bifunctional 5'-3' exo/endoribonuclease RNase J. *Struct. Lond. Engl. 1993* **19**, 1252–1261 (2011).

188. Li de la Sierra-Gallay, I., Zig, L., Jamalli, A. & Putzer, H. Structural insights into the dual activity of RNase J. *Nat. Struct. Mol. Biol.* **15**, 206–212 (2008).

189. Deikus, G., Condon, C. & Bechhofer, D. H. Role of *Bacillus subtilis* RNase J1 endonuclease and 5'-exonuclease activities in *trp* leader RNA turnover. *J. Biol. Chem.* **283**, 17158–17167 (2008).

190. Condon, C. What is the role of RNase J in mRNA turnover? *RNA Biol.* **7**, 316–321 (2010).

191. Even, S. *et al.* Ribonucleases J1 and J2: two novel endoribonucleases in *B. subtilis* with functional homology to *E. coli* RNase E. *Nucleic Acids Res.* **33**, 2141–2152 (2005).

192. Mathy, N. *et al.* Bacillus subtilis ribonucleases J1 and J2 form a complex with altered enzyme behaviour. *Mol. Microbiol.* **75**, 489–498 (2010).

193. Linder, P., Lemeille, S. & Redder, P. Transcriptome-wide analyses of 5'-ends in RNase J mutants of a gram-positive pathogen reveal a role in RNA maturation, regulation and degradation. *PLoS Genet.* **10**, e1004207 (2014).

194. Liu, N. *et al.* The *Streptococcus mutans irvA* gene encodes a trans-acting riboregulatory mRNA. *Mol. Cell* **57**, 179–190 (2015).

195. Bugrysheva, J. V. & Scott, J. R. The ribonucleases J1 and J2 are essential for growth and have independent roles in mRNA decay in *Streptococcus pyogenes*. *Mol. Microbiol.* **75**, 731–743 (2010).

196. Grunberg-Manago, M. & Ochoa, S. Enzymatic synthesis and breakdown of polynucleotides; polynucleotide phosphorylase1. *J. Am. Chem. Soc.* **77**, 3165–3166 (1955).

197. Grunberg-Manago, M., Ortiz, P. J. & Ochoa, S. Enzymatic synthesis of nucleic acidlike polynucleotides. *Science* **122**, 907–910 (1955).

198. Mohanty, B. K. & Kushner, S. R. Polynucleotide phosphorylase functions both as a 3' → 5' exonuclease and a poly(A) polymerase in *Escherichia coli*. *Proc. Natl. Acad. Sci.* **97**, 11966–11971 (2000).

199. Symmons, M. F., Jones, G. H. & Luisi, B. F. A duplicated fold is the structural basis for polynucleotide phosphorylase catalytic activity, processivity, and regulation. *Struct. Lond. Engl. 1993* **8**, 1215–1226 (2000).

200. Fazal, F. M., Koslover, D. J., Luisi, B. F. & Block, S. M. Direct observation of processive exoribonuclease motion using optical tweezers. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 15101–15106 (2015).

201. Wang, W. & Bechhofer, D. H. Properties of a *Bacillus subtilis* polynucleotide phosphorylase deletion strain. *J. Bacteriol.* **178**, 2375–2382 (1996).

202. Bechhofer, D. H. & Stasinopoulos, S. J. *tetA*(L) mutants of a tetracycline-sensitive strain of *Bacillus subtilis* with the polynucleotide phosphorylase gene deleted. *J. Bacteriol.* **180**, 3470–3473 (1998).

203. Luttinger, A., Hahn, J. & Dubnau, D. Polynucleotide phosphorylase is necessary for competence development in *Bacillus subtilis*. *Mol. Microbiol.* **19**, 343–356 (1996).

204. Lécrivain, A.-L. *et al. In vivo* 3'-to-5' exoribonuclease targetomes of *Streptococcus pyogenes*. *Proc. Natl. Acad. Sci.* **115**, 11814–11819 (2018).

205. De Lay, N. & Gottesman, S. Role of polynucleotide phosphorylase in sRNA function in *Escherichia coli*. *RNA* **17**, 1172–1189 (2011).

206. Sesto, N. *et al.* A PNPase dependent CRISPR system in *Listeria*. *PLoS Genet.* **10**, e1004065 (2014).

207. Awano, N. *et al. Escherichia coli* RNase R has dual activities, helicase and RNase. *J. Bacteriol.* **192**, 1344–1352 (2010).

208. Hossain, S. T., Malhotra, A. & Deutscher, M. P. The helicase activity of ribonuclease R is essential for efficient nuclease activity. *J. Biol. Chem.* **290**, 15697–15706 (2015).

209. Oussenko, I. A. & Bechhofer, D. H. The *yvaJ* gene of *Bacillus subtilis* encodes a 3'-to-5' exoribonuclease and is not essential in a strain lacking polynucleotide phosphorylase. *J. Bacteriol.* **182**, 2639–2642 (2000).

210. Sulthana, S. & Deutscher, M. P. Multiple exoribonucleases catalyze maturation of the 3' terminus of 16S ribosomal RNA (rRNA). *J. Biol. Chem.* **288**, 12574–12579 (2013).

211. Bernstein, J. A., Lin, P.-H., Cohen, S. N. & Lin-Chao, S. Global analysis of *Escherichia coli* RNA degradosome function using DNA microarrays. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 2758–2763 (2004).

212. Chen, C. & Deutscher, M. P. RNase R is a highly unstable protein regulated by growth phase and stress. *RNA* **16**, 667–672 (2010).

213. Cairrão, F., Cruz, A., Mori, H. & Arraiano, C. M. Cold shock induction of RNase R and its role in the maturation of the quality control mediator SsrA/tmRNA. *Mol. Microbiol.* **50**, 1349–1360 (2003).

214. Zhang, Y. *et al.* A Stress response that monitors and regulates mRNA structure is central to cold shock adaptation. *Mol. Cell* **70**, 274-286.e7 (2018).

215. Oussenko, I. A., Sanchez, R. & Bechhofer, D. H. *Bacillus subtilis* YhaM, a member of a new family of 3'-to-5' exonucleases in Gram-positive bacteria. *J. Bacteriol.* **184**, 6250–6259 (2002).

216. Zhang, Q., Soares de Oliveira, S., Colangeli, R. & Gennaro, M. L. Binding of a novel host factor to the pT181 plasmid replication enhancer. *J. Bacteriol.* **179**, 684–688 (1997).

217. Au, N. *et al.* Genetic composition of the *Bacillus subtilis* SOS System. *J. Bacteriol.* **187**, 7655–7666 (2005).

218. Mechold, U., Fang, G., Ngo, S., Ogryzko, V. & Danchin, A. YtqI from *Bacillus subtilis* has both oligoribonuclease and pAp-phosphatase activity. *Nucleic Acids Res.* **35**, 4552–4561 (2007).

219. Schmier, B. J., Nelersa, C. M. & Malhotra, A. Structural basis for the bidirectional activity of *Bacillus* nanoRNase NrnA. *Sci. Rep.* **7**, 11085 (2017).

220. Bandyra, K. J. & Luisi, B. F. RNase E and the high-fidelity orchestration of RNA metabolism. *Microbiol. Spectr.* **6**, (2018).

221. Tejada-Arranz, A., de Crécy-Lagard, V. & de Reuse, H. Bacterial RNA degradosomes: molecular machines under tight control. *Trends Biochem. Sci.* (2019). Preprint at: https://doi.org/10.1016/j.tibs.2019.10.002.

222. Redder, P. Molecular and genetic interactions of the RNA degradation machineries in Firmicute bacteria. *Wiley Interdiscip. Rev. RNA* **9**, e1460 (2018).

223. Cho, K. H. The Structure and function of the Gram-positive bacterial RNA degradosome. *Front. Microbiol.* **8**, 154 (2017).

224. Roux, C. M., DeMuth, J. P. & Dunman, P. M. Characterization of components of the *Staphylococcus aureus* mRNA degradosome holoenzyme-like complex. *J. Bacteriol.* **193**, 5520–5526 (2011).

225. Giraud, C. *et al.* The C-terminal region of the RNA helicase CshA is required for the interaction with the degradosome and turnover of bulk RNA in the opportunistic pathogen *Staphylococcus aureus*. *RNA Biol.* **12**, 658–674 (2015).

226. Mu, R., Shinde, P., Zou, Z., Kreth, J. & Merritt, J. Examining the protein interactome and subcellular localization of RNase J2 complexes in *Streptococcus mutans*. *Front. Microbiol.* **10**, 2150 (2019).

227. Salvo, E., Alabi, S., Liu, B., Schlessinger, A. & Bechhofer, D. H. Interaction of *Bacillus subtilis* polynucleotide phosphorylase and RNase Y: structural mapping and effect on mRNA turnover. *J. Biol. Chem.* **291**, 6655–6663 (2016).

228. Morita, T., Kawamoto, H., Mizota, T., Inada, T. & Aiba, H. Enolase in the RNA degradosome plays a crucial role in the rapid decay of glucose transporter mRNA in the response to phosphosugar stress in *Escherichia coli*. *Mol. Microbiol.* **54**, 1063–1075 (2004).

229. Cascante-Estepa, N., Gunka, K. & Stülke, J. Localization of components of the RNA-degrading machine in *Bacillus subtilis*. *Front. Microbiol.* **7**, 1492 (2016).

230. Hadjeras, L. *et al.* Detachment of the RNA degradosome from the inner membrane of *Escherichia coli* results in a global slowdown of mRNA degradation, proteolysis of RNase E and increased turnover of ribosome-free transcripts. *Mol. Microbiol.* **111**, 1715–1731 (2019).

231. Brantl, S. & Brückner, R. Small regulatory RNAs from low-GC Gram-positive bacteria. *RNA Biol.* **11**, 443–456 (2014).

232. Obana, N., Shirahama, Y., Abe, K. & Nakamura, K. Stabilization of *Clostridium perfringens* collagenase mRNA by VR-RNA-dependent cleavage in 5' leader sequence. *Mol. Microbiol.* **77**, 1416–1428 (2010).

233. Ramirez-Peña, E., Treviño, J., Liu, Z., Perez, N. & Sumby, P. The group A Streptococcus small regulatory RNA FasX enhances streptokinase activity by increasing the stability of the *ska* mRNA transcript. *Mol. Microbiol.* **78**, 1332–1347 (2010).

234. Deutscher, M. P. How bacterial cells keep ribonucleases under control. *FEMS Microbiol. Rev.* **39**, 350–361 (2015).

235. Bardwell, J. C. *et al.* Autoregulation of RNase III operon by mRNA processing. *EMBO J.* **8**, 3401–3407 (1989).

236. Mayer, J. E. & Schweiger, M. RNase III is positively regulated by T7 protein kinase. *J. Biol. Chem.* **258**, 5340–5343 (1983).

237. Kim, K., Manasherob, R. & Cohen, S. N. YmdB: a stress-responsive ribonuclease-binding regulator of *E. coli* RNase III activity. *Genes Dev.* **22**, 3497–3508 (2008).

238. Koch, G. *et al.* Attenuating *Staphylococcus aureus* virulence by targeting flotillin protein scaffold activity. *Cell Chem. Biol.* **24**, 845-857.e6 (2017).

239. Jamalli, A., Hébert, A., Zig, L. & Putzer, H. Control of expression of the RNases J1 and J2 in *Bacillus subtilis*. *J. Bacteriol.* **196**, 318–324 (2014).

240. Jarrige, A.-C., Mathy, N. & Portier, C. PNPase autocontrols its expression by degrading a double-stranded structure in the *pnp* mRNA leader. *EMBO J.* **20**, 6845–6855 (2001).

241. Carzaniga, T., Dehò, G. & Briani, F. RNase III-independent autogenous regulation of *Escherichia coli* polynucleotide phosphorylase via translational repression. *J. Bacteriol.* **197**, 1931–1938 (2015).

242. Park, H., Yakhnin, H., Connolly, M., Romeo, T. & Babitzke, P. CsrA participates in a PNPase autoregulatory mechanism by selectively repressing translation of *pnp* transcripts that have been previously processed by RNase III and PNPase. *J. Bacteriol.* **197**, 3751–3759 (2015).

243. Fontaine, F. *et al.* The small RNA SraG participates in PNPase homeostasis. *RNA* **22**, 1560–1573 (2016).

244. Zangrossi, S. *et al.* Transcriptional and post-transcriptional control of polynucleotide phosphorylase during cold acclimation in *Escherichia coli*. *Mol. Microbiol.* **36**, 1470–1480 (2000).

245. Mathy, N., Jarrige, A.-C., Robert-Le Meur, M. & Portier, C. Increased expression of *Escherichia coli* polynucleotide phosphorylase at low temperatures is linked to a decrease in the efficiency of autocontrol. *J. Bacteriol.* **183**, 3848–3854 (2001).

246. Nurmohamed, S. *et al.* Polynucleotide phosphorylase activity may be modulated by metabolites in *Escherichia coli*. *J. Biol. Chem.* **286**, 14315–14323 (2011).

247. Siculella, L. *et al.* Guanosine 5'-diphosphate 3'-diphosphate (ppGpp) as a negative modulator of polynucleotide phosphorylase activity in a 'rare' actinomycete. *Mol. Microbiol.* **77**, 716–729 (2010).

248. Del Favero, M. *et al.* Regulation of *Escherichia coli* polynucleotide phosphorylase by ATP. *J. Biol. Chem.* **283**, 27355–27359 (2008).

249. Liang, W. & Deutscher, M. P. Transfer-messenger RNA-SmpB protein regulates ribonuclease R turnover by promoting binding of HslUV and Lon Proteases. *J. Biol. Chem.* **287**, 33472–33479 (2012).

250. Liang, W., Malhotra, A. & Deutscher, M. P. Acetylation regulates the stability of a bacterial protein: growth stage-dependent modification of RNase R. *Mol. Cell* **44**, 160–166 (2011).

251. Liang, W. & Deutscher, M. P. Ribosomes regulate the stability and action of the exoribonuclease RNase R. *J. Biol. Chem.* **288**, 34791–34798 (2013).

252. Chen, R. *et al.* Polynucleotide phosphorylase regulates multiple virulence factors and the stabilities of small RNAs RsmY/Z in *Pseudomonas aeruginosa*. *Front. Microbiol.* **7**, 247 (2016).

253. Redko, Y. *et al.* RNase J depletion leads to massive changes in mRNA abundance in *Helicobacter pylori*. *RNA Biol.* **13**, 243–253 (2016).

254. Pobre, V. & Arraiano, C. M. Next generation sequencing analysis reveals that the ribonucleases RNase II, RNase R and PNPase affect bacterial motility and biofilm formation in *E. coli*. *BMC Genomics* **16**, 72 (2015).

255. Leskinen, K., Varjosalo, M. & Skurnik, M. Absence of YbeY RNase compromises the growth and enhances the virulence plasmid gene expression of *Yersinia enterocolitica* O:3. *Microbiol. Read. Engl.* **161**, 285–299 (2015).

256. Chen, H., Shiroguchi, K., Ge, H. & Xie, X. S. Genome-wide study of mRNA degradation and transcript elongation in *Escherichia coli*. *Mol. Syst. Biol.* **11**, 781 (2015).

257. Kristoffersen, S. M. *et al.* Global mRNA decay analysis at single nucleotide resolution reveals segmental and positional degradation patterns in a Gram-positive bacterium. *Genome Biol.* **13**, R30 (2012).

258. Waters, S. A. *et al.* Small RNA interactome of pathogenic *E. coli* revealed through crosslinking of RNase E. *EMBO J.* **36**, 374–387 (2017).

259. Lybecker, M., Zimmermann, B., Bilusic, I., Tukhtubaeva, N. & Schroeder, R. The double-stranded transcriptome of *Escherichia coli*. *Proc. Natl. Acad. Sci.* **111**, 3134–3139 (2014).

260. Gordon, G. C., Cameron, J. C. & Pfleger, B. F. RNA Sequencing identifies new RNase III cleavage sites in *Escherichia coli* and reveals increased regulation of mRNA. *mBio* **8**, (2017).

261. Altuvia, Y. *et al. In vivo* cleavage rules and target repertoire of RNase III in *Escherichia coli*. *Nucleic Acids Res.* **46**, 10380–10394 (2018).

262. Cooper, D. A., Jha, B. K., Silverman, R. H., Hesselberth, J. R. & Barton, D. J. Ribonuclease L and metal-ion-independent endoribonuclease cleavage sites in host and viral RNAs. *Nucleic Acids Res.* **42**, 5202–5216 (2014).

263. Dar, D. *et al.* Term-seq reveals abundant ribo-regulation of antibiotics resistance in bacteria. *Science* **352**, aad9822 (2016).

264. Dar, D. & Sorek, R. High-resolution RNA 3'-ends mapping of bacterial Rho-dependent transcripts. *Nucleic Acids Res.* **46**, 6797–6805 (2018).

265. Dar, D. & Sorek, R. Extensive reshaping of bacterial operons by programmed mRNA decay. *PLoS Genet.* **14**, e1007354 (2018).

266. Beall, B., Facklam, R. & Thompson, T. Sequencing emm-specific PCR products for routine and accurate typing of group A streptococci. *J. Clin. Microbiol.* **34**, 953–958 (1996).

267. Sanderson-Smith, M. *et al.* A systematic and functional classification of *Streptococcus pyogenes* that serves as a new tool for molecular typing and vaccine development. *J. Infect. Dis.* **210**, 1325–1338 (2014).

268. Walker, M. J. *et al.* Disease manifestations and pathogenic mechanisms of Group A Streptococcus. *Clin. Microbiol. Rev.* **27**, 264–301 (2014).

269. Wessels, M. R. Pharyngitis and Scarlet Fever. in *Streptococcus pyogenes: Basic Biology to Clinical Manifestations* (eds. Ferretti, J. J., Stevens, D. L. & Fischetti, V. A.) (University of Oklahoma Health Sciences Center, 2016).

270. Stevens, D. L. & Bryant, A. E. Severe Group A Streptococcal Infections. in *Streptococcus pyogenes: Basic Biology to Clinical Manifestations* (eds. Ferretti, J. J., Stevens, D. L. & Fischetti, V. A.) (University of Oklahoma Health Sciences Center, 2016).

271. Efstratiou, A. & Lamagni, T. Epidemiology of *Streptococcus pyogenes*. in *Streptococcus pyogenes: basic biology to clinical manifestations* (eds. Ferretti, J. J., Stevens, D. L. & Fischetti, V. A.) (University of Oklahoma Health Sciences Center, 2016).

272. Lamagni, T. L. *et al.* Epidemiology of severe *Streptococcus pyogenes* disease in Europe. *J. Clin. Microbiol.* **46**, 2359–2367 (2008).

273. O'Grady, K.-A. F. *et al.* The epidemiology of invasive group A streptococcal disease in Victoria, Australia. *Med. J. Aust.* **186**, 565–569 (2007).

274. O'Loughlin, R. E. *et al.* The epidemiology of invasive group A streptococcal infection and potential vaccine implications: United States, 2000-2004. *Clin. Infect. Dis. Off. Publ. Infect. Dis. Soc. Am.* **45**, 853–862 (2007).

275. Cunningham, M. W. Post-Streptococcal autoimmune sequelae: rheumatic fever and beyond. in *Streptococcus pyogenes : Basic Biology to Clinical Manifestations* (eds. Ferretti, J. J., Stevens, D. L. & Fischetti, V. A.) (University of Oklahoma Health Sciences Center, 2016).

276. Carapetis, J. R., Steer, A. C., Mulholland, E. K. & Weber, M. The global burden of group A streptococcal diseases. *Lancet Infect. Dis.* **5**, 685–694 (2005).

277. Cattoir, V. Mechanisms of antibiotic resistance. in *Streptococcus pyogenes: basic biology to clinical manifestations* (eds. Ferretti, J. J., Stevens, D. L. & Fischetti, V. A.) (University of Oklahoma Health Sciences Center, 2016).

278. Spellerberg, B. & Brandt, C. Laboratory diagnosis of *Streptococcus pyogenes* (group A streptococci). in *Streptococcus pyogenes: basic biology to clinical manifestations* (eds. Ferretti, J. J., Stevens, D. L. & Fischetti, V. A.) (University of Oklahoma Health Sciences Center, 2016).

279. Dale, J. B. *et al.* Group A streptococcal vaccines: paving a path for accelerated development. *Vaccine* **31**, B216–B222 (2013).

280. Dale, J. B. *et al.* Current approaches to Group A Streptococcal vaccine development. in *Streptococcus pyogenes: basic biology to clinical manifestations* (eds. Ferretti, J. J., Stevens, D. L. & Fischetti, V. A.) (University of Oklahoma Health Sciences Center, 2016).

281. Nobbs, A. H., Lamont, R. J. & Jenkinson, H. F. *Streptococcus* adherence and colonization. *Microbiol. Mol. Biol. Rev.* **73**, 407–450 (2009).

282. Ralph, A. P. & Carapetis, J. R. Group A streptococcal diseases and their global burden. in *Host-Pathogen Interactions in Streptococcal Diseases* (ed. Chhatwal, G. S.) 1–27 (Springer Berlin Heidelberg, 2012).

283. Cole, J. N., Barnett, T. C., Nizet, V. & Walker, M. J. Molecular insight into invasive group A streptococcal disease. *Nat. Rev. Microbiol.* **9**, 724–736 (2011).

284. Johansson, L., Thulin, P., Low, D. E. & Norrby-Teglund, A. Getting under the skin: the immunopathogenesis of *Streptococcus pyogenes* deep tissue infections. *Clin. Infect. Dis.* **51**, 58–65 (2010).

285. Bisno, A., Brito, M. & Collins, C. Molecular basis of group A streptococcal virulence. *Lancet Infect. Dis.* **3**, 191–200 (2003).

286. Fraser, J. D. & Proft, T. The bacterial superantigen and superantigen-like proteins. *Immunol. Rev.* **225**, 226–243 (2008).

287. Hynes, W. & Sloan, M. Secreted extracellular virulence factors. in *Streptococcus pyogenes : Basic Biology to Clinical Manifestations* (eds. Ferretti, J. J., Stevens, D. L. & Fischetti, V. A.) (University of Oklahoma Health Sciences Center, 2016).

288. Chaussee, M. S., Phillips, E. R. & Ferretti, J. J. Temporal production of streptococcal erythrogenic toxin B (streptococcal cysteine proteinase) in response to nutrient depletion. *Infect. Immun.* **65**, 1956–1959 (1997).

289. Yu, C. E. & Ferretti, J. J. Frequency of the erythrogenic toxin B and C genes (*speB* and *speC*) among clinical isolates of group A streptococci. *Infect. Immun.* **59**, 211–215 (1991).

290. Carroll, R. K. & Musser, J. M. From transcription to activation: how group A streptococcus, the flesh-eating pathogen, regulates SpeB cysteine protease production. *Mol. Microbiol.* **81**, 588–601 (2011).

291. Liu, T.-Y. & Elliott, S. D. Streptococcal Proteinase: the zymogen to enzyme transformation. *J. Biol. Chem.* **240**, 1138–1142 (1965).

292. Olsen, J. G., Dagil, R., Niclasen, L. M., Sørensen, O. E. & Kragelund, B. B. Structure of the mature Streptococcal cysteine protease exotoxin mSpeB in its active dimeric form. *J. Mol. Biol.* **393**, 693–703 (2009).

293. Chen, C.-Y. *et al.* Maturation processing and characterization of Streptopain. *J. Biol. Chem.* **278**, 17336–17343 (2003).

294. Nelson, D. C., Garbe, J. & Collin, M. Cysteine proteinase SpeB from *Streptococcus pyogenes* – a potent modifier of immunologically important host and bacterial proteins. *Biol. Chem.* **392**, 1077–1088 (2011).

295. Kapur, V., Majesky, M. W., Li, L. L., Black, R. A. & Musser, J. M. Cleavage of interleukin 1 beta (IL-1 beta) precursor to produce active IL-1 beta by a conserved extracellular cysteine protease from *Streptococcus pyogenes*. *Proc. Natl. Acad. Sci.* **90**, 7676–7680 (1993).

296. LaRock, C. N. *et al.* IL-1β is an innate immune sensor of microbial proteolysis. *Sci. Immunol.* **1**, (2016).

297. Sumby, P., Whitney, A. R., Graviss, E. A., DeLeo, F. R. & Musser, J. M. Genome-wide analysis of group a streptococci reveals a mutation that modulates global phenotype and disease specificity. *PLoS Pathog.* **2**, e5 (2006).

298. Cole, J. N. *et al.* Trigger for group A streptococcal M1T1 invasive disease. *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.* **20**, 1745–1747 (2006).

299. Aziz, R. K. *et al.* Invasive M1T1 group A Streptococcus undergoes a phase-shift *in vivo* to prevent proteolytic degradation of multiple virulence factors by SpeB. *Mol. Microbiol.* **51**, 123–134 (2004).

300. Olsen, R. J. *et al.* The majority of 9,729 Group A streptococcus strains causing disease secrete SpeB cysteine protease: pathogenesis implications. *Infect. Immun.* **83**, 4750–4758 (2015).

301. Neely, M. N., Lyon, W. R., Runft, D. L. & Caparon, M. Role of RopB in growth phase expression of the SpeB cysteine protease of *Streptococcus pyogenes*. *J. Bacteriol.* **185**, 5166–5174 (2003).

302. Loughman, J. A. & Caparon, M. Regulation of SpeB in *Streptococcus pyogenes* by pH and NaCl: a model for *in vivo* gene expression. *J. Bacteriol.* **188**, 399–408 (2006).

303. Lyon, W. R., Gibson, C. M. & Caparon, M. G. A role for trigger factor and an Rgg-like regulator in the transcription, secretion and processing of the cysteine proteinase of *Streptococcus pyogenes*. *EMBO J.* **17**, 6263–6275 (1998).

304. Chaussee, M. S., Ajdic, D. & Ferretti, J. J. The *rgg* gene of *Streptococcus pyogenes* NZ131 positively influences extracellular SpeB production. *Infect. Immun.* **67**, 1715–1722 (1999).

305. Anbalagan, S., Dmitriev, A., McShan, W. M., Dunman, P. M. & Chaussee, M. S. Growth phase-dependent modulation of Rgg binding specificity in *Streptococcus pyogenes*. *J. Bacteriol.* **194**, 3961–3971 (2012).

306. Federle, M. J., McIver, K. S. & Scott, J. R. A response regulator that represses transcription of several virulence operons in the group A streptococcus. *J. Bacteriol.* **181**, 3649–3657 (1999).

307. Heath, A., DiRita, V. J., Barg, N. L. & Engleberg, N. C. A two-component regulatory system, CsrR-CsrS, represses expression of three *Streptococcus pyogenes* virulence factors, hyaluronic acid capsule, streptolysin S, and pyrogenic exotoxin B. *Infect. Immun.* **67**, 5298–5305 (1999).

308. Miller, A. A., Engleberg, N. C. & DiRita, V. J. Repression of virulence genes by phosphorylation-dependent oligomerization of CsrR at target promoters in *S. pyogenes*. *Mol. Microbiol.* **40**, 976–990 (2001).

309. Treviño, J. *et al.* CovS simultaneously activates and inhibits the CovR-mediated repression of distinct subsets of group A Streptococcus virulence factor-encoding genes. *Infect. Immun.* **77**, 3141–3149 (2009).

310. Malke, H., Steiner, K., McShan, W. M. & Ferretti, J. J. Linking the nutritional status of *Streptococcus pyogenes* to alteration of transcriptional gene expression: the action of CodY and RelA. *Int. J. Med. Microbiol. IJMM* **296**, 259–275 (2006).

311. Loughman, J. A. & Caparon, M. G. A novel adaptation of aldolase regulates virulence in *Streptococcus pyogenes*. *EMBO J.* **25**, 5414–5422 (2006).

312. Ma, Y., Bryant, A. E., Salmi, D. B., McIndoo, E. & Stevens, D. L. *vfr*, a novel locus affecting cysteine protease production in *Streptococcus pyogenes*. *J. Bacteriol.* **191**, 3189–3194 (2009).

313. Shelburne, S. A. *et al.* An N-terminal signal peptide of Vfr protein negatively influences RopB-dependent SpeB expression and attenuates virulence in *Streptococcus pyogenes*. *Mol. Microbiol.* **82**, 1481–1495 (2011).

314. Makthal, N. *et al.* Structural and functional analysis of RopB: a major virulence regulator in *Streptococcus pyogenes*. *Mol. Microbiol.* **99**, 1119–1133 (2016).

315. Do, H. *et al.* Leaderless secreted peptide signaling molecule alters global gene expression and increases virulence of a human bacterial pathogen. *Proc. Natl. Acad. Sci. U. S. A.* **114**, E8498–E8507 (2017).

316. Brouwer, S. *et al.* The endopeptidase PepO regulates the SpeB cysteine protease and is essential for the virulence of invasive M1T1 *Streptococcus pyogenes*. *J. Bacteriol.* **200**, e00654–17 (2018).

317. Vega, L. A., Malke, H. & McIver, K. S. Virulence-related transcriptional regulators of *Streptococcus pyogenes*. in *Streptococcus pyogenes: basic biology to clinical manifestations* (eds. Ferretti, J. J., Stevens, D. L. & Fischetti, V. A.) (University of Oklahoma Health Sciences Center, 2016).

318. McDowell, E. J., Callegari, E. A., Malke, H. & Chaussee, M. S. CodY-mediated regulation of *Streptococcus pyogenes* exoproteins. *BMC Microbiol.* **12**, 114 (2012).

319. Ribardo, D. A. & McIver, K. S. Defining the Mga regulon: Comparative transcriptome analysis reveals both direct and indirect regulation by Mga in the group A streptococcus. *Mol. Microbiol.* **62**, 491–508 (2006).

320. Marouni, M. J. & Sela, S. The *luxS* gene of *Streptococcus pyogenes* regulates expression of genes that affect internalization by epithelial cells. *Infect. Immun.* **71**, 5633–5639 (2003).

321. Reid, S. D. *et al.* Inactivation of the group A Streptococcus regulator *srv* results in chromosome wide reduction of transcript levels, and changes in extracellular levels of Sic and SpeB. *FEMS Immunol. Med. Microbiol.* **48**, 283–292 (2006).

322. Marouni, M. J., Ziomek, E. & Sela, S. Influence of group A streptococcal acid glycoprotein on expression of major virulence factors and internalization by epithelial cells. *Microb. Pathog.* **35**, 63–72 (2003).

323. Molinari, G. *et al.* The role played by the group A streptococcal negative regulator Nra on bacterial interactions with epithelial cells. *Mol. Microbiol.* **40**, 99–114 (2001).

324. Churchward, G. The two faces of Janus: virulence gene regulation by CovR/S in group A streptococci. *Mol. Microbiol.* **64**, 34–41 (2007).

325. Kietzman, C. C. & Caparon, M. G. CcpA and LacD.1 affect temporal regulation of *Streptococcus pyogenes* virulence genes. *Infect. Immun.* **78**, 241–252 (2010).

326. Shelburne, S. A. *et al.* A combination of independent transcriptional regulators shapes bacterial virulence gene expression during infection. *PLoS Pathog.* **6**, e1000817 (2010).

327. Li, Z., Sledjeski, D. D., Kreikemeyer, B., Podbielski, A. & Boyle, M. D. P. Identification of *pel*, a *Streptococcus pyogenes* locus that affects both surface and secreted proteins. *J. Bacteriol.* **181**, 6019–6027 (1999).

328. Mangold, M. *et al.* Synthesis of group A streptococcal virulence factors is controlled by a regulatory RNA molecule. *Mol. Microbiol.* **53**, 1515–1527 (2004).

329. Chen, Z., Itzek, A., Malke, H., Ferretti, J. J. & Kreth, J. Dynamics of *speB* mRNA transcripts in *Streptococcus pyogenes*. *J. Bacteriol.* **194**, 1417–1426 (2012).

330. Lyon, W. R. & Caparon, M. G. Trigger factor-mediated prolyl isomerization influences maturation of the *Streptococcus pyogenes* cysteine protease. *J. Bacteriol.* **185**, 3661–3667 (2003).

331. Ma, Y. *et al.* Identification and characterization of bicistronic *speB* and *prsA* gene expression in the Group A Streptococcus. *J. Bacteriol.* **188**, 7626–7634 (2006).

332. Kagawa, T. F., O'toole, P. W. & Cooney, J. C. SpeB-Spi: a novel protease-inhibitor pair from *Streptococcus pyogenes*. *Mol. Microbiol.* **57**, 650–666 (2005).

333. Kreikemeyer, B., McIver, K. S. & Podbielski, A. Virulence factor regulation and regulatory networks in *Streptococcus pyogenes* and their impact on pathogen-host interactions. *Trends Microbiol.* **11**, 224–232 (2003).

334. Tart, A. H., Walker, M. J. & Musser, J. M. New understanding of the group A Streptococcus pathogenesis cycle. *Trends Microbiol.* **15**, 318–325 (2007).

335. Graham, M. R. *et al.* Virulence control in group A Streptococcus by a two-component gene regulatory system: global expression profiling and in vivo infection modeling. *Proc. Natl. Acad. Sci.* **99**, 13855–13860 (2002).

336. Opdyke, J. A., Scott, J. R. & Moran, Jr., C. P. Expression of the secondary sigma factor σX in *Streptococcus pyogenes* is restricted at two levels. *J. Bacteriol.* **185**, 4291–4297 (2003).

337. Woodbury, R. L., Wang, X. & Moran, C. P. Sigma X induces competence gene expression in *Streptococcus pyogenes*. *Res. Microbiol.* **157**, 851–856 (2006).

338. Jimenez, J. C. & Federle, M. J. Quorum sensing in group A Streptococcus. *Front. Cell. Infect. Microbiol.* **4**, (2014).

339. Liu, Z., Treviño, J., Ramirez-Peña, E. & Sumby, P. The small regulatory RNA FasX controls pilus expression and adherence in the human bacterial pathogen Group A streptococcus. *Mol. Microbiol.* **86**, 140–154 (2012).

340. Danger, J. L. *et al.* The small regulatory RNA FasX enhances Group A streptococcus virulence and inhibits pilus expression via serotype-specific targets. *Mol. Microbiol.* **96**, 249–262 (2015).

341. Danger, J. L., Makthal, N., Kumaraswami, M. & Sumby, P. The FasX small regulatory RNA negatively regulates the expression of two fibronectin-binding proteins in Group A streptococcus. *J. Bacteriol.* **197**, 3720–3730 (2015).

342. Barnett, T. C., Bugrysheva, J. V. & Scott, J. R. Role of mRNA stability in growth phase regulation of gene expression in the Group A streptococcus. *J. Bacteriol.* **189**, 1866–1873 (2007).

343. Bugrysheva, J. V. & Scott, J. R. Regulation of virulence gene expression in *Streptococcus pyogenes*: determinants of differential mRNA decay. *RNA Biol.* **7**, 569–572 (2010).

344. Broglia, L. *et al.* RNase Y-mediated regulation of the streptococcal pyrogenic exotoxin B. *RNA Biol.* **15**, 1336–1347 (2018).

345. Lambert, J. M., Bongers, R. S. & Kleerebezem, M. Cre-lox-based system for multiple gene deletions and selectable-marker removal in Lactobacillus plantarum. *Appl. Environ. Microbiol.* **73**, 1126–1135 (2007).

346. Majumder, H. K., Maitra, U. & Rosenberg, M. Termination of transcription by bacteriophage T3 RNA polymerase: homogeneous 3'-terminal oligonucleotide sequence of *in vitro* T3 RNA polymerase transcripts. *Proc. Natl. Acad. Sci. U. S. A.* **76**, 5110–5113 (1979).

347. Geissendörfer, M. & Hillen, W. Regulated expression of heterologous genes in *Bacillus subtilis* using the Tn10 encoded tet regulatory elements. *Appl. Microbiol. Biotechnol.* **33**, 657–663 (1990).

348. Wang, B. & Kuramitsu, H. K. Inducible antisense RNA expression in the characterization of gene functions in *Streptococcus mutans*. *Infect. Immun.* **73**, 3568–3576 (2005).

349. Sambrook, J., Fritsch, E. F. & Maniatis, T. *Molecular cloning: a laboratory manual*, New York. Cold Spring Harbor Laboratory Press (1989).

350. Sitkiewicz, I. & Musser, J. M. Expression microarray and mouse virulence analysis of four conserved two-component gene regulatory systems in Group A streptococcus. *Infect. Immun.* **74**, 1339–1351 (2006).

351. Yanisch-Perron, C., Vieira, J. & Messing, J. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33**, 103–119 (1985).

352. Resch, U. *et al.* A two-component regulatory system impacts extracellular membrane-derived vesicle production in Group A streptococcus. *mBio* **7**, (2016).

353. Siller, M. *et al.* Functional analysis of the group A streptococcal luxS/AI-2 system in metabolism, adaptation to stress and interaction with host cells. *BMC Microbiol.* **8**, 188 (2008).

354. Charpentier, E. *et al.* Novel cassette-based shuttle vector system for gram-positive bacteria. *Appl. Environ. Microbiol.* **70**, 6076–6085 (2004).

355. Silva-Rocha, R. *et al.* The standard european vector architecture (SEVA): a coherent platform for the analysis and deployment of complex prokaryotic phenotypes. *Nucleic Acids Res.* **41**, D666–D675 (2013).

356. Que, Y. A., Haefliger, J. A., Francioli, P. & Moreillon, P. Expression of *Staphylococcus aureus* clumping factor A in *Lactococcus lactis* subsp.*cremoris* using a new shuttle vector. *Infect. Immun.* **68**, 3516–3522 (2000).

357. Loh, J. M. S. & Proft, T. Toxin–antitoxin-stabilized reporter plasmids for biophotonic imaging of Group A streptococcus. *Appl. Microbiol. Biotechnol.* **97**, 9737–9745 (2013).

358. Wang, W. & Malcolm, B. A. Two-stage PCR protocol allowing introduction of multiple mutations, deletions and insertions using QuikChange site-directed mutagenesis. *BioTechniques* **26**, 680–682 (1999).

359. Le Rhun, A., Beer, Y. Y., Reimegård, J., Chylinski, K. & Charpentier, E. RNA sequencing uncovers antisense RNAs and novel small RNAs in *Streptococcus pyogenes*. *RNA Biol.* **13**, 177–195 (2016).

360. Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).

361. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**, 10–12 (2011).

362. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinforma. Oxf. Engl.* **29**, 15–21 (2013).

363. Ferretti, J. J. *et al.* Complete genome sequence of an M1 strain of *Streptococcus pyogenes*. *Proc. Natl. Acad. Sci.* **98**, 4658–4663 (2001).

364. Robinson, J. T. *et al.* Integrative genomics viewer. *Nat. Biotechnol.* **29**, 24–26 (2011).

365. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).

366. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinforma. Oxf. Engl.* **26**, 139–140 (2010).

367. McCarthy, D. J., Chen, Y. & Smyth, G. K. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res.* **40**, 4288–4297 (2012).

368. Crooks, G. E., Hon, G., Chandonia, J.-M. & Brenner, S. E. WebLogo: a sequence logo generator. *Genome Res.* **14**, 1188–1190 (2004).

369. Lorenz, R. *et al.* ViennaRNA Package 2.0. *Algorithms Mol. Biol. AMB* **6**, 26 (2011).

370. Darty, K., Denise, A. & Ponty, Y. VARNA: interactive drawing and editing of the RNA secondary structure. *Bioinformatics* **25**, 1974–1975 (2009).

371. Chen, Z., Raghavan, R., Qi, F., Merritt, J. & Kreth, J. Genome-wide screening of potential RNase Y-processed mRNAs in the M49 serotype *Streptococcus pyogenes* NZ131. *MicrobiologyOpen* **8**, e00671 (2018).

372. Levin, J. C. & Wessels, M. R. Identification of *csrR/csrS*, a genetic locus that regulates hyaluronic acid capsule synthesis in group A Streptococcus. *Mol. Microbiol.* **30**, 209–219 (1998).

373. Casinhas, J., Matos, R. G., Haddad, N. & Arraiano, C. M. Biochemical characterization of *Campylobacter jejuni* PNPase, an exoribonuclease important for bacterial pathogenicity. *Biochimie* **147**, 70–79 (2018).

374. Spickler, C. & Mackie, G. A. Action of RNase II and polynucleotide phosphorylase against RNAs containing stem-loops of defined structure. *J. Bacteriol.* **182**, 2422–2427 (2000).

375. Li, Z. & Deutscher, M. P. Maturation pathways for *E. coli* tRNA precursors: a random multienzyme process in vivo. *Cell* **86**, 503–512 (1996).

376. Liu, B., Kearns, D. B. & Bechhofer, D. H. Expression of multiple *Bacillus subtilis* genes is controlled by decay of *slrA* mRNA from Rho-dependent 3' ends. *Nucleic Acids Res.* **44**, 3364–3372 (2016).

377. Rivas, E., Klein, R. J., Jones, T. A. & Eddy, S. R. Computational identification of noncoding RNAs in *E. coli* by comparative genomics. *Curr. Biol.* **11**, 1369–1373 (2001).

378. Aseev, L. V., Levandovskaya, A. A., Tchufistova, L. S., Scaptsova, N. V. & Boni, I. V. A new regulatory circuit in ribosomal protein operons: S2-mediated control of the *rpsB-tsf* expression *in vivo*. *RNA* **14**, 1882–1894 (2008).

379. Nasser, W. *et al.* Evolutionary pathway to increased virulence and epidemic Group A streptococcus disease derived from 3,615 genome sequences. *Proc. Natl. Acad. Sci. U. S. A.* **111**, E1768–E1776 (2014).

380. Rasmussen, M., Müller, H. P. & Björck, L. Protein GRAB of *Streptococcus pyogenes* regulates proteolysis at the bacterial surface by binding alpha2-macroglobulin. *J. Biol. Chem.* **274**, 15336–15344 (1999).

381. Nyberg, P., Rasmussen, M. & Björck, L. alpha2-Macroglobulin-proteinase complexes protect *Streptococcus pyogenes* from killing by the antimicrobial peptide LL-37. *J. Biol. Chem.* **279**, 52820–52823 (2004).

382. Grunert, H.-P. *et al.* Studies on RNase T1 mutants affecting enzyme catalysis. *Eur. J. Biochem.* **197**, 203–207 (1991).

383. Yoshida, H. The Ribonuclease T1 family. in *Methods in Enzymology* (ed. Nicholson, A. W.) vol. **341** 28–41 (Academic Press, 2001).

384. Mora, L., Ngo, S., Laalami, S. & Putzer, H. Chapter Fourteen - *In vitro* study of the major *Bacillus subtilis* ribonucleases Y and J. in *Methods in Enzymology* (ed. Carpousis, A. J.) vol. **612** 343–359 (Academic Press, 2018).

385. Deikus, G., Babitzke, P. & Bechhofer, D. H. Recycling of a regulatory protein by degradation of the RNA to which it binds. *Proc. Natl. Acad. Sci.* **101**, 2747–2751 (2004).

386. Meinken, C., Blencke, H.-M., Ludwig, H. & Stülke, J. Expression of the glycolytic *gapA* operon in *Bacillus subtilis*: differential syntheses of proteins encoded by the operon. *Microbiol. Read. Engl.* **149**, 751–761 (2003).

387. Gerwig, J. & Stülke, J. Caught in the act: RNA-Seq provides novel insights into mRNA degradation. *Mol. Microbiol.* **94**, 5–8 (2014).

Appendix

6.1 Supplementary tables

Gene ID	Gene name	Log 2 FC	FDR
SPy_0033	*purE*	-1.26	2.80E-02
Spy_sRNA112285		-1.60	4.03E-03
SPy_0167	*slo*	-1.44	3.41E-02
SPy_0183	*opuAA*	1.20	1.32E-02
SPy_0212	*speG*	1.30	1.54E-02
SPy_0374	*ribU*	1.71	2.80E-02
SPy_0384	*fhuB.1*	-1.36	1.02E-02
SPy_0405		1.08	2.55E-02
Spy_sRNA350644		-2.02	3.49E-03
SPy_0460	*rplK*	1.28	1.97E-02
SPy_0461	*rplA*	1.41	9.32E-03
Spy_sRNA480642		-2.15	1.49E-03
SPy_0636	*idnO*	-2.28	2.14E-03
SPy_0658		2.57	1.52E-03
SPy_0659		3.37	5.24E-03
SPy_0660		2.98	4.25E-02
SPy_0661		3.43	7.57E-03
SPy_0663		3.30	3.99E-04
SPy_0664		2.62	4.47E-04
SPy_0665		3.03	2.31E-03
SPy_0666		2.91	5.57E-04
SPy_0667		2.55	5.57E-04
SPy_0669		2.64	3.99E-04
SPy_0670		2.55	2.14E-03
SPy_0671		2.48	5.45E-04
SPy_0672		2.60	1.24E-02
SPy_0673		2.79	3.39E-03
SPy_0676		2.08	3.81E-03
SPy_0679		1.74	1.17E-02
SPy_0681		1.79	1.26E-02
SPy_0686		1.78	8.29E-03
SPy_0688		1.56	9.32E-03
SPy_0689		1.51	3.04E-02
SPy_0693		1.72	4.04E-02
SPy_0697		1.72	4.86E-03
SPy_0698		2.12	3.49E-03
SPy_0701	*hylP1*	1.93	5.24E-03

SPy_0702		1.96	3.39E-03
SPy_0705		2.04	2.01E-02
SPy_0707		2.20	1.45E-02
SPy_0710		1.79	3.49E-03
SPy_0723		-1.19	2.26E-02
Spy_sRNA579683	*rplS* leader	2.08	1.65E-03
SPy_0724	*rplS*	3.10	3.99E-04
SPy_0888	*clpL*	1.22	1.60E-02
SPy_1097	*folE*	1.28	7.38E-03
SPy_1098	*folP*	1.64	3.25E-03
SPy_1099	*folQ*	1.75	3.25E-03
SPy_1100	*folK*	1.76	7.89E-03
SPy_1270		-1.34	1.90E-02
Spy_sRNA1109453		-2.06	3.07E-02
SPy_1357	*grab*	1.40	8.60E-03
SPy_1470		2.09	5.40E-03
Spy_sRNA1260092		-1.87	3.49E-03
SPy_1539	*asnA*	-1.22	1.90E-02
SPy_1633	*rny*	-4.36	1.91E-04
SPy_1643	*mapZ*	1.61	4.86E-03
SPy_1644	*ypsC*	2.29	5.27E-04
SPy_1646	*gpsB*	2.59	5.57E-04
SPy_1823	*comEB*	-1.21	3.04E-02
SPy_1842	*spi*	1.11	3.04E-02
SPy_1871	*rpsN2*	1.95	3.39E-03
SPy_1894	*pyrG*	1.27	2.02E-02
SPy_2016	*sic*	-1.70	1.02E-02
SPy_2038		-3.76	1.91E-04
SPy_2039	*speB*	-3.56	5.27E-04
Spy_sRNA1699993		-4.02	3.99E-04
SPy_2040		-4.12	1.91E-04
SPy_2043	*mf*	-1.13	2.00E-02
SPy_2070	*groEL*	1.97	2.75E-03
SPy_2072	*groES*	2.04	1.66E-03
SPy_2092	*rpsB*	3.44	5.03E-04
SPy_2093	*tsf*	3.08	5.01E-04
SPy_2126		1.94	8.29E-03
SPy_2127		2.02	1.60E-02
SPy_2128		2.85	7.38E-03
SPy_2173		-2.04	9.44E-04
SPy_2174		-1.22	3.04E-02
SPy_2188	*mnmA*	1.30	8.42E-03
SPy_2191		1.08	2.00E-02

Table A1. Genes differentially expressed in the Δ*rny* strain compared to the WT strain.
The genes that are encoded within the same transcript are clustered in a grey box. The genes encoded in the SF370.1 prophage are indicated in grey. For each gene, the information is provided with regard the gene ID, the gene name, the log 2 Fold Change (Log 2 FC) (in red downregulated and in green upregulated genes) and the False Discovery Rate (FDR). R.A.-B. generated the raw list of differentially expressed genes analysed by L.B, A.-L.L. and A.L.R.

Position	Strand	nt in 3′	Log 2 FC	FDR	Gene ID	Description	Location	Type
336198	+	G	-3.9122153	0.0080557	*SPy_0397* ↔ *tmk*	"hypothetical protein" ↔ "thymidylate kinase"	UTR	U
480622	+	G	-4.5503693	0.00117632	*SPy_0596* ↔ Spy_sRNA480642	"hypothetical protein"	UTR	U
1152937	−	G	-2.3191867	0.01106928	Spy_sRNA1153040		UTR	S
1382711	−	C	-4.260725	0.00242634	*pbpX*	penicillin binding protein 2X	ORF	U
1763486	−	G	-4.6695369	0.00045468	*SPy_2111*	hypothetical protein	ORF	U
625092	+	G	-3.7094857	0.00474057	*pheS* ↔ *pheT*	"phenylalanyl-tRNA synthetase subunit alpha" ↔ "phenylalanyl-tRNA synthetase subunit beta"	UTR	U
112341	−	G	-4.9238213	0.00041443	*SPy_0121* ↔ Spy_sRNA112285	"deoxyguanosine kinase/deoxyadenosine kinase(I) subunit"	UTR	U
193042	+	G	-3.5806271	0.00154055	*pgi* ↔ *rivR*	"glucose-6-phosphate isomerase" ↔ "regulatory protein"	UTR	S
275237	+	G	-3.6375752	0.00201823	*SPy_0309*	hypothetical protein	ORF	U
302413	+	T	-3.806466	0.00275994	*murC*	UDP-N-acetylmuramate--L-alanine ligase	ORF	U
473749	+	G	-4.6483965	0.00352477	*SPy_0587*	hypothetical protein	ORF	U
1401255	−	G	-3.7594454	0.00320088	*glyS*	glycyl-tRNA synthetase subunit beta	ORF	U
744449	+	G	-3.3234456	0.00433312	*SPy_0899* ↔ Spy_sRNA744487	"hypothetical protein"	UTR	U
323093	+	G	-4.2776828	0.00158438	*hlyX*	hemolysin	ORF	U
1014753	−	G	-3.0954977	0.02448428	*cdd* ↔ *bmpA*	"cytidine deaminase" ↔ "lipoprotein"	UTR	U
737938	+	G	-4.1187432	0.00477073	Spy_sRNA737945		UTR	U
1711879	−	G	-3.9218442	0.0047934	*SPy_2054*	transcriptional regulator	ORF	U
1276533	−	G	-2.9368277	0.03815276	*SPy_1551*	hypothetical protein	ORF	U
409647	−	G	-3.5102573	0.00133109	*SPy_0508*	hypothetical protein	ORF	U
480754	−	G	-4.5455981	0.00073023	*gpmB1*	phosphoglycerate mutase	ORF	U
1300735	−	G	-3.7014986	0.00414061	*SPy_1580*	acetate kinase	ORF	U
318951	+	G	-3.4146157	0.0029544	*trmL*	23S rRNA methyltransferase	ORF	U
287642	−	T	-4.7302255	0.00074043	*ntpJ*	V-type Na -ATPase subunit J	ORF	U
1761437	−	G	-3.7753131	0.00120921	*nrdD*	anaerobic ribonucleoside triphosphate reductase	ORF	S
662728	+	G	-5.1033771	0.00117632	*rplT* ↔ *ltaS*	"50S ribosomal protein L20" ↔ "hypothetical protein"	UTR	U
1403423	−	G	-3.996955	0.00081057	*glyQ* ↔ *glyS*	"glycyl-tRNA synthetase subunit alpha" ↔ "glycyl-tRNA synthetase subunit beta"	UTR	U

1430377	−	G	-2.8108126	0.00891298	*trmB* ↔ *SPy_t51*	"tRNA (guanine-N(7)-)-methyltransferase" ↔ "tRNA-Ser (GGA)"	UTR	S
838265	+	G	-3.4485711	0.00665644	*acoC* ↔ *acoL*	"branched-chain alpha-keto acid dehydrogenase E2 subunit" ↔ "dihydrolipoamide dehydrogenase%2C component E3"	UTR	U
1634045	−	A	-4.0575361	0.00491335	*proS*	prolyl-tRNA synthetase	ORF	U
1152840	−	G	-3.6947204	0.00472336	Spy_sRNA1153040		UTR	U
1353157	−	G	-5.2863681	0.00030048	*rpoZ*	DNA-directed RNA polymerase subunit omega	ORF	S
1571171	−	G	-3.6304043	0.00143485	*rpoE* ↔ Spy_sRNA1571135	"DNA-directed RNA polymerase subunit delta"	UTR	U
499822	+	A	-4.9577492	0.00117632	*SPy_0622*	hypothetical protein	ORF	U
1796021	−	A	-3.283546	0.04036051	*aspS*	aspartyl-tRNA synthetase	ORF	U
1109507	−	G	-3.889542	0.00081057	*SPy_1340*	hypothetical protein	ORF	U
1625568	−	A	-2.3886405	0.0074938	*rpsO*	30S ribosomal protein S15	ORF	U
1766433	−	G	-4.691884	0.000803	*recA*	recombinase A	ORF	U
1558584	−	G	-3.1672063	0.00688493	*glnA*	glutamine synthetase	ORF	U
1400285	−	G	-3.1794718	0.00073023	Spy_sRNA1400292		UTR	U
1225472	−	G	-4.3622941	0.00037248	*SPy_1494* ↔ *SPy_1493*	"hypothetical protein" ↔ "hypothetical protein"	UTR	S
381164	+	G	-3.5383645	0.00154055	*frr* ↔ *SPy_0464*	"ribosome recycling factor" ↔ "hypothetical protein"	UTR	U
1057924	−	G	-4.5936208	0.0023665	*pfkA* ↔ *pyk*	"6-phosphofructokinase" ↔ "pyruvate kinase"	UTR	U
72783	+	A	-2.8157838	0.00511808	*rpsE*	30S ribosomal protein S5	ORF	U
472165	+	G	-3.668581	0.00205669	*ptsK*	HPr kinase/phosphorylase	ORF	U
1247251	−	G	-3.6161234	0.00339062	*sepF*	hypothetical protein	ORF	U
1023988	−	G	-3.3367194	0.02200802	*pstB*	phosphate transporter ATP-binding protein	ORF	U
1764034	−	G	-3.9296616	0.03322003	*SPy_2111*	hypothetical protein	ORF	U
1611283	−	G	-3.4501462	0.00426525	*SPy_1936*	hypothetical protein	ORF	U
1345770	−	G	-2.3184931	0.01759037	*stkP*	protein kinase	ORF	U
305395	+	G	-2.7598824	0.01365923	*SPy_0348*	aminodeoxychorismate lyase	ORF	U
713911	+	G	-4.0989164	0.00154055	*SPy_0865* ↔ Spy_sRNA713953	"hypothetical protein"	UTR	U
518762	+	G	-2.9654672	0.0040871	*ftsX*	cell-division protein	ORF	U
947461	+	G	-3.0050101	0.00805848	*gyrA*	DNA gyrase subunit A	ORF	U
1530955	+	G	-2.5164073	0.01134356	*SPy_1845*	hypothetical protein	ORF	U

71140	+	G	-4.2553806	0.00162751	*rpsH*	30S ribosomal protein S8	ORF	U
665855	+	G	-3.5871165	0.00255899	*rlmI*	hypothetical protein	ORF	U
614992	+	G	-4.1151462	0.00317484	*atpB*	ATP synthase F0F1 subunit A	ORF	U
1085336	−	G	-3.4725101	0.00869713	*SPy_1308* ↔ Spy_sRNA1085297	"esterase"	UTR	U
1850570	+	G	-5.2963182	0.00117632	*htrA*	serine protease	ORF	U
107969	−	G	-4.3375232	0.00180923	*proC* ↔ *SPy_0110*	"pyrroline-5-carboxylate reductase" ↔ "hypothetical protein"	UTR	U
1517566	+	G	-4.2630377	0.00187064	*corA*	divalent cation transport protein	ORF	U
320197	+	G	-4.4041088	0.02266403	*SPy_0374*	hypothetical protein	ORF	U
1568240	−	G	-3.6909479	0.0036685	*SPy_1892* ↔ *SPy_t52*	"hypothetical protein" ↔ "tRNA-Leu (AAG)"	UTR	U
115048	−	G	-3.6502608	0.00046297	*rofA*	regulatory protein	ORF	U
1424789	−	G	-3.6244504	0.00414061	*infB*	translation initiation factor IF-2	ORF	U
290364	+	G	-4.3589183	0.00300344	*htpX*	heat shock protein HtpX	ORF	U
88476	+	G	-6.1135341	0.00154055	*adcB*	ABC transporter permease	ORF	S
1403406	−	G	-4.3945381	0.00046297	*glyQ* ↔ *glyS*	"glycyl-tRNA synthetase subunit alpha" ↔ "glycyl-tRNA synthetase subunit beta"	UTR	S
120684	+	G	-3.2637025	0.00303308	*SPy_0129*	hypothetical protein	ORF	S
1717865	+	G	-2.4389867	0.01191043	*SPy_2063*	hypothetical protein	ORF	S
1454339	−	G	-3.2705283	0.00266228	*fabH*	3-oxoacyl-ACP synthase	ORF	S
734526	+	G	-4.5910271	0.00778311	*SPy_0887*	hypothetical protein	ORF	U
1450213	−	A	-2.2433414	0.04441892	*fabF*	3-oxoacyl-ACP synthase	ORF	U
320563	+	G	-4.0009619	0.00158438	*SPy_0374*	hypothetical protein	ORF	U
71149	+	G	-5.898757	0.00154055	*rpsH* ↔ *rplF*	"30S ribosomal protein S8" ↔ "50S ribosomal protein L6"	UTR	U
1453867	−	G	-3.4262488	0.00762705	*fabH*	3-oxoacyl-ACP synthase	ORF	U
1838142	−	G	-3.5203423	0.00251456	*glcU*	hypothetical protein	ORF	U
111220	+	G	-5.2044057	0.00162196	*ytpR*		ORF	U
1329354	+	G	-2.6896869	0.02268772	*SPy_1608*	hypothetical protein	ORF	U
1680611	−	G	-6.7989094	0.00030048	Spy_sRNA1680670		UTR	U
1183494	−	G	-3.8298352	0.00274189	*niaX*	hypothetical protein	ORF	U
428218	+	G	-4.0545526	0.00382502	*vicK*	two-component sensor histidine kinase	ORF	U

472174	+	G	-4.4255648	0.00154055	*ptsK*	HPr kinase/phosphorylase	ORF	U
32908	+	G	-4.1072367	0.00207313	*prsA*	ribose-phosphate pyrophosphokinase	ORF	U
254023	−	G	-3.291518	0.01753913	*dacA_2*		ORF	U
112328	−	G	-3.9559779	0.00235458	*SPy_0121* ↔ Spy_sRNA112285	"deoxyguanosine kinase/deoxyadenosine kinase(I) subunit"	UTR	U
849003	+	G	-3.2023381	0.00717095	*glmM*	phosphoglucosamine mutase	ORF	U
1472043	−	G	-3.4971512	0.01513906	*SPy_1781* ↔ *SPy_1780*	"hypothetical protein" ↔ "hypothetical protein"	UTR	U
1699224	−	C	-6.054102	0.00037248	*speB*	pyrogenic exotoxin B	ORF	U
146735	+	A	-2.3924809	0.03999118	*purA*	adenylosuccinate synthetase	ORF	U
1406622	−	G	-3.1630325	0.02349569	*nagA*	N-acetylglucosamine-6-phosphate deacetylase	ORF	U
471430	+	G	-4.8222547	0.00117632	Spy_sRNA471434		UTR	U
248410	+	G	-4.8186438	0.00171405	*rgpG*	hypothetical protein	ORF	U
1701817	−	G	-4.1326919	0.00688493	*mf* ↔ *ropB*	"mitogenic factor" ↔ "transcriptional regulator"	UTR	U
274907	+	G	-3.5356776	0.02853138	*SPy_0309*	hypothetical protein	ORF	U
1366777	+	G	-4.418621	0.00154055	*recU*	Holliday junction-specific endonuclease	ORF	U
978434	−	G	-3.2392114	0.01513906	*citC*	citrate lyase synthetase	ORF	U
1535210	−	G	-4.5322096	0.00073023	*SPy_1850* ↔ *pfl*	"esterase" ↔ "pyruvate formate-lyase"	UTR	U
1375132	−	G	-3.533737	0.01201342	*aapA*	amino acid permease	ORF	U
1571150	−	G	-4.3566986	0.00074043	*rpoE* ↔ Spy_sRNA1571135	"DNA-directed RNA polymerase subunit delta"	UTR	U
1345722	−	G	-3.2121506	0.00168462	*stkP*	protein kinase	ORF	U
115985	−	G	-3.1139618	0.00973486	*rofA*	regulatory protein	ORF	U
1383473	−	G	-4.5363271	0.00266228	*ftsL*	cell division protein	ORF	U
73054	+	G	-4.0339542	0.00167333	*rpmD* ↔ Spy_sRNA73113	"50S ribosomal protein L30"	UTR	U
1840448	−	G	-3.5148879	0.01902795	*guaB*	inosine 5'-monophosphate dehydrogenase	ORF	U
307062	−	G	-3.4535335	0.00205495	*yidC2* ↔ Spy_sRNA307010	"OxaA-like protein"	UTR	S
1699189	−	C	-5.7379285	0.00074224	*speB*	pyrogenic exotoxin B	ORF	U
381172	+	G	-3.3579912	0.00255899	*frr* ↔ *SPy_0464*	"ribosome recycling factor" ↔ "hypothetical protein"	UTR	U
1152301	−	G	-3.3321089	0.0274186	*alaS*	alanyl-tRNA synthetase	ORF	U
477779	−	G	-2.9536274	0.00727361	Spy_sRNA477984 ↔ Spy_sRNA477714		UTR	U

285038	+	G	-5.3092415	0.00479199	*sstT*	serine/threonine transporter SstT	ORF	U
1683210	−	G	-4.0481137	0.0027707	*sic*	inhibitor of complement-mediated lysis	ORF	S
849024	+	G	-3.1773496	0.00528794	*glmM*	phosphoglucosamine mutase	ORF	U
319062	+	G	-4.5428746	0.00154055	*trmL*	23S rRNA methyltransferase	ORF	U
932728	+	G	-4.6742514	0.00154055	*pbuX*	purine permease	ORF	U
278753	+	C	-3.6248272	0.00334461	*SPy_0316*	hypothetical protein	ORF	U
1264364	−	G	-5.6432185	0.00074043	*SPy_1536* ↔ *lacI*	"hypothetical protein" ↔ "ribose transport operon repressor"	UTR	U
1139348	−	G	-1.8191738	0.04036051	*pstI*	phosphoenolpyruvate:sugar phosphotransferase system enzyme I	ORF	U
1680618	−	G	-4.5139229	0.00321851	Spy_sRNA1680670		UTR	U
65047	+	C	-2.152743	0.03204506	*rplD*	50S ribosomal protein L4	ORF	U
784641	−	G	-2.4638029	0.01615945	*SPy_0948 or SPy_0947*	"hypothetical protein" or "hypothetical protein"	ORF	U
1518318	+	G	-2.5784585	0.02063761	*SPy_1828*	hypothetical protein	ORF	U
183554	−	G	-2.9797315	0.00278041	*SPy_0201*	hypothetical protein	ORF	U
1535306	−	G	-3.500358	0.0074938	*SPy_1850*	esterase	ORF	U
1792766	−	G	-4.2382581	0.01671002	*SPy_2153*	hypothetical protein	ORF	U
115218	−	G	-2.9362253	0.04248731	*rofA*	regulatory protein	ORF	U
407561	−	G	-2.5414217	0.01328359	*pcp*	pyrrolidone-carboxylate peptidase	ORF	U
308001	−	G	-3.6760674	0.00956642	*yidC2*	OxaA-like protein	ORF	U
1439397	+	G	-2.7989535	0.01109809	*SPy_1737*	hypothetical protein	ORF	U
1404769	−	G	-2.3931691	0.01108144	Spy_sRNA1404921	*glyQ* leader	UTR	S
705947	+	G	-4.2332012	0.00324595	*fruB*	fructose-1-phosphate kinase	ORF	U
18649	+	G	-3.3719277	0.00304669	*SPy_r01* ↔ *SPy_t01*	"16S rRNA"	UTR	U
14744	+	G	-2.1411051	0.04359308	*ftsH*	cell division protein	ORF	U
958809	+	G	-4.1543631	0.00154055	*topA*	DNA topoisomerase I	ORF	U
1792652	+	G	-5.1825736	0.00717095	*SPy_2152*	hypothetical protein	ORF	U
1581533	−	G	-2.7451403	0.01102107	*SPy_r12* ↔ *SPy_t55*	"16S rRNA" ↔ "tRNA-Ala (TGC)"	UTR	U
78458	+	G	-3.5684251	0.00158438	*rplQ*	50S ribosomal protein L17	ORF	U
24650	+	G	-2.9598792	0.00753443	*SPy_r03* ↔ *SPy_t11*	"16S rRNA"	UTR	U

640306	+	C	-2.8899624	0.00675131	*dnaG*	DNA primase	ORF	S
335393	+	G	-4.9382046	0.00117632	Spy_sRNA335396		UTR	U
265934	+	G	-2.7358625	0.00964512	*SPy_r07* ↔ *SPy_t35*	"16S rRNA" ↔ "tRNA-Ala (TGC)"	UTR	U
287634	−	G	-4.5329337	0.01118653	*ntpJ*	V-type Na -ATPase subunit J	ORF	U
1611380	−	G	-3.679853	0.01378971	*SPy_1936*	hypothetical protein	ORF	U
1387143	−	G	-4.1460129	0.00232382	*SPy_1673*	ABC transporter permease	ORF	U
85408	+	G	-2.5395422	0.00831754	*SPy_t34*	tRNA-Leu (CAA)	UTR	U
1698455	−	C	-5.5444909	0.00074043	*SPy_2038*		ORF	U
80868	+	G	-2.9256586	0.00954855	*SPy_r05* ↔ *SPy_t22*	"16S rRNA" ↔ "tRNA-Ala (TGC)"	UTR	U
1030602	−	G	-3.4153439	0.00517723	*spxA*	transcriptional regulator Spx	ORF	U
1683123	−	G	-4.1484974	0.00163858	*sic*	inhibitor of complement-mediated lysis	ORF	S
1455284	−	G	-4.4045308	0.00545934	*phaB*	enoyl-CoA hydratase	ORF	U
1053328	−	A	-3.5205225	0.02585314	*phnA*	hypothetical protein	ORF	U
1247216	−	G	-4.8773062	0.0038952	*sepF*	hypothetical protein	ORF	U
311860	+	G	-3.9528814	0.00562052	*SPy_0358*	hypothetical protein	ORF	U
1762537	−	G	-4.6397886	0.00037248	*SPy_2111* ↔ *nrdD*	"hypothetical protein" ↔ "anaerobic ribonucleoside triphosphate reductase"	UTR	U
1557813	+	G	-3.9401261	0.00402886	*rnjA* ↔ *glnA*	"Zn-dependent hydrolase" ↔ "glutamine synthetase"	UTR	U
110128	−	G	-3.7188665	0.00242634	*SPy_0116*	hypothetical protein	ORF	U
1441577	+	G	-2.9233771	0.00964512	*manM*	PTS system mannose-specific transporter subunit IIC	ORF	U
122080	+	G	-2.3953095	0.02741635	*SPy_0130* ↔ *SPy_0131*	"hypothetical protein" ↔ "hypothetical protein"	UTR	S
897129	+	G	-3.9037623	0.00229742	*folC.1*	folyl-polyglutamate synthetase	ORF	U
1844913	+	G	-2.5923234	0.01624666	*SPy_2210*	ABC transporter ATP-binding protein	ORF	U
1540175	−	G	-4.0135435	0.00604315	*norA*	antibiotic resistance protein NorA	ORF	S
131575	−	G	-5.3159644	0.00522435	*SPy_0144*	hypothetical protein	ORF	U
1454152	−	G	-3.7756169	0.02327768	*fabH*	3-oxoacyl-ACP synthase	ORF	U
1713230	−	G	-3.1481282	0.00419439	*secE* ↔ *pflA*	"preprotein translocase subunit SecE" ↔ "pyruvate formate-lyase activating protein"	UTR	U
1764633	−	G	-3.2913278	0.01671002	*SPy_2112*	hypothetical protein	ORF	U
1685503	−	A	-2.5441635	0.01722349	*mga*	M protein trans-acting positive regulator	ORF	U

31422	+	G	-2.9299432	0.00493291	*cdhA*	hypothetical protein	ORF	U
29085	+	A	-4.1223532	0.00117632	*SPy_t21* ↔ *cdhA*	"tRNA-Ser (GCT)" ↔ "hypothetical protein"	UTR	S
227166	+	G	-2.8702295	0.0104923	*ksgA* ↔ *rsgA*	"dimethyladenosine transferase" ↔ "ribosome-associated GTPase"	UTR	U
1390062	−	G	-2.9645959	0.02154857	*tkt*	transketolase	ORF	U
1345714	−	G	-2.6791934	0.00891298	*stkP*	protein kinase	ORF	U
307075	−	G	-3.8061917	0.01009089	*yidC2* ↔ Spy_sRNA307010	"OxaA-like protein"	UTR	U
370709	−	G	-3.5303362	0.03857277	*mntR*	metal-dependent transcriptional regulator	ORF	U
1329406	+	G	-2.7201455	0.03299831	*SPy_1608* ↔ Spy_sRNA1329429	"hypothetical protein"	UTR	U
93367	+	G	-3.0547674	0.00239447	Spy_sRNA93359		UTR	U
335380	+	G	-3.0857496	0.00451315	*clpP*	ATP-dependent Clp protease proteolytic subunit	ORF	U
1699202	−	T	-5.3650616	0.00503233	*speB*	pyrogenic exotoxin B	ORF	U
1698333	−	T	-6.1964279	0.00081057	*SPy_2038* ↔ *prsA2*	"foldase PrsA"	UTR	U
75005	+	G	-3.835325	0.00154055	*secY*	preprotein translocase subunit SecY	ORF	U
1055635	−	G	-3.1799145	0.00565414	*sipC*	signal peptidase I	ORF	U
1683190	−	G	-3.4714639	0.0141975	*sic*	inhibitor of complement-mediated lysis	ORF	U
1697886	−	G	-4.3049012	0.00688493	*SPy_2038* ↔ *prsA2*	"foldase PrsA"	UTR	U
69330	+	A	-2.7252457	0.0183345	*rplN*	50S ribosomal protein L14	ORF	U
117541	+	G	-3.2802471	0.00875631	*cpa*		ORF	U
1616523	−	G	-4.0750371	0.03226786	*cysE* ↔ *SPy_1942*	"serine acetyltransferase" ↔ "hypothetical protein"	UTR	U
1698650	−	T	-5.2403324	0.00073023	*SPy_2038*		ORF	U
292611	+	G	-3.0832761	0.00344565	*covS*	sensory transduction histidine kinase	ORF	U
177858	+	G	-2.7200762	0.02015808	*SPy_0190 or SPy_0191a*	"hypothetical protein" or "hypothetical protein"	ORF	U
1141021	−	G	-2.7372841	0.00522435	*ptsH*	phosphocarrier protein HPr	ORF	U
1698505	−	T	-7.1420783	0.00352409	*SPy_2038*		ORF	U

Table A2. RNA 5' ends more abundant in the WT strain than in Δ*rny* strain (5' *rny*_ends).
List of the RNA 5' ends identified as more abundant in the WT strain than in the Δ*rny* strain and named as 5' *rny*_ends (see Figure 11). The information is provided with regard the chromosome position, the directionality of the strand (positive (+) or negative (-)), the log 2 Fold Change (Log2FC), the False Discovery Rate (FDR), the gene ID (the symbol "↔"

depicts the intergenic region between two genes), the description of the gene function based on NCBI, the location of the 5' *rny*_end (open reading frame (ORF) and untranslated region (UTR)) and the type of 5' *rny*_end ("unique" (U) or "stepped" (S)). R.A.-B. generated the raw list of 5' *rny*_ends analysed by L.B, A.-L.L. and A.L.R.

Clones	Sequence	Position 5′ end	Position 3′ end
1	TAGATTAACTTATTGTCTATTACC	− 137 nt	− 25 nt
2	−	n.d.	n.d.
3	−	n.d.	n.d.
4	−	n.d.	n.d.
5	−	n.d.	n.d.
6	GTATAAAAAAAATGATTATTGTCTATTACC	− 137 nt	− 10 nt
7	AAAGAAATTAGGTGTTCTATTACCATTCAT	− 131 nt	+ 23 nt
8	AAATGCAGTAGATTATTATTGTCTATTACC	− 137 nt	− 27 nt
9	AAATGCAGTAGATTATTATTGTCTATTACC	− 137 nt	− 27 nt
10	CAAATGCAGTAGATTTTATTGTCTATTACC	− 137 nt	− 28 nt
11a	GCAGTAGATTATTATTGTCTATT	− 137 nt	− 27 nt
11b	ATGCAGTAGATTAACTATTGTCTATTACCA	− 136 nt	− 25 nt
12	AAATGCAGTAGATTATTATTGTCTATTACC	− 137 nt	− 27 nt
13	−	n.d	n.d
14	GTAGATTTTATTGTCTATTACC	− 137 nt	− 28 nt
15a	AATGCAGTAGATTAGTTATTGTCTATTACC	− 137 nt	− 28 nt
15b	AAATGCAGTAGATTAGTTATTGTCTATTAC	− 138 nt	− 27 nt
16	AATTAGGTTATTGTCTATTACC	− 137 nt	+ 20 nt
17	−	n.d.	n.d.
18	AAATGCAGTAGATTATCTATTACCATTCAT	− 131	− 26 nt
19	−	n.d.	n.d.
20	AAATGCAGTAGATTATTATTGTCTATTACC	− 137 nt	− 27 nt
21a	CAAATGCAGTAGATTTTATTGTCTATTAC	− 137 nt	− 28 nt
21b	CAAATGCAGTAGATTTTATTGTCTATTACC	− 137 nt	− 28 nt
22a	AATGCAGTAGATTAATTATTGTCTATTACC	− 137 nt	− 26 nt
22b	CAAATGCAGTAGATTTCTATTACCATTCAT	− 131 nt	− 28 nt
23	AATGCAGTAGATTAATATTGTCTATTACCA	− 136 nt	− 26 nt
24	AAATGCAGTAGATTATTATTGTCTATTACC	− 137 nt	− 27 nt
25	CAAATGCAGTAGATTTTATTGTCTATTACC	− 137 nt	− 28 nt
26	ATGCAGTAGATTAACTTATTGTCTATTACC	− 137 nt	− 25 nt
27	−	n.d.	n.d.
28	CAAATGCAGTAGATTTTATTGTCTATTACCA	− 137 nt	− 28 nt
29	AATGCAGTAGATTAATTATTGTCTATTACC	− 137 nt	− 26 nt
30	AAATGCAGTAGATTATTATTGTCTATTACC	− 137 nt	− 27 nt
31	CAAATGCAGTAGATTTCTATTACCATTCAT	− 131 nt	− 28 nt
32	TAGATTAATTATTGTCTATTACC	− 137 nt	− 26 nt

Table A3. Simultaneous mapping of the *speB* sRNA 5' and 3' ends.
List of the detected *speB*-sRNA 5' (in purple) and 3' (in green) ends obtained by circularization of RNA by self-ligation followed by Reverse Transcriptase-Polymerase Chain Reaction (Circ-RT-PCR; see Material & Methods and Figure 17). The DNA sequences amplified from the cDNA were cloned into the TOPO TA vector (Invitrogen) and analyzed by Sanger sequencing. The positions reported for the RNA 5' and 3' ends are relative to the *speB* start codon (see Figure 17). The n.d. (not determined) indicates the clones for which the transcript 5' and 3' could not be retrieved either because the DNA sequencing signal was undetectable or because the sequenced vector did not contain any inserts. For the clones 11, 15, 21 and 22 two Circ-RT-

PCR generated inserts were found in the vector (indicated as “a” and “b”) and both sequences were included in the analysis.

Position	Strand	nt in 3'	Log 2 FC	FDR	Gene ID	Description	Location	type
1763385	−	T	-3.8568929	5.00E-05	*SPy_2111*	hypothetical protein	ORF	S
477545	−	T	-3.7580326	3.36E-05	*SPy_0593* (antisense)	hypothetical protein	UTR	S
861256	+	T	-3.072974	0.00852775	Spy_sRNA860730 ↔ *lepA*	CRISPR RNA (crRNA) ↔ "GTP-binding protein LepA"	UTR	U
1152709	−	G	-3.6611274	0.00040452	Spy_sRNA1153040 ↔ *alaS*	"alanyl-tRNA synthetase"	UTR	U
292690	+	C	-2.0623033	0.00809917	*covS*	sensory transduction histidine kinase	ORF	S
417307	+	A	-3.3282281	0.00768874	Spy_sRNA416968 ↔ *thrS*	"threonyl-tRNA synthetase"	UTR	S
377298	+	C	-3.0978246	0.0038518	*ftsK*	hypothetical protein	ORF	S
287540	−	A	-2.8618134	0.00863008	*ntpJ*	V-type Na -ATPase subunit J	ORF	S
1633958	−	T	-4.1623222	0.0005108	*proS*	prolyl-tRNA synthetase	ORF	U
1152119	−	A	-3.1195535	0.00113354	*alaS*	alanyl-tRNA synthetase	ORF	S
656504	+	T	-4.1491298	0.01898369	*SPy_0797*	hypothetical protein	ORF	U
1516070	−	A	-5.6467011	0.00073311	*uvrA*	excinuclease ABC subunit A	ORF	U
1405169	−	G	-4.7914094	0.00019191	*SPy_1691*	hypothetical protein	ORF	U
524653	+	T	-2.4776189	0.02886727	*aspC*	aspartate aminotransferase	ORF	U
111312	+	T	-5.0038511	0.00409174	*ytpR*		ORF	S
1401103	−	T	-3.7757093	0.00254882	*glyS* ↔ *SPy_1687*	"glycyl-tRNA synthetase subunit beta" ↔ "hypothetical protein"	UTR	U
350705	+	A	-3.7833051	0.00260679	Spy_sRNA350644		UTR	U
381252	+	A	-4.1598184	0.04574839	*frr* ↔ *SPy_0464*	"ribosome recycling factor" ↔ "hypothetical protein"	UTR	U
323184	+	G	-4.1824682	0.00375107	*hlyX*	hemolysin	ORF	U
1023862	−	T	-3.1170428	0.01256691	*pstB*	phosphate transporter ATP-binding protein	ORF	U
1762451	−	A	-2.7001608	0.00059059	*nrdD*	anaerobic ribonucleoside triphosphate reductase	ORF	S
117634	+	C	-2.3210181	0.00863928	*cpa*		ORF	S
1811851	−	C	-2.9226296	0.00098742	Spy_sRNA1811977 ↔ *rpsD*	"30S ribosomal protein S4"	UTR	U
583794	+	T	-4.0273433	0.00316825	*gyrB* ↔ *ezrA*	"DNA gyrase subunit B" ↔ "septation ring formation regulator EzrA"	UTR	U
1276420	−	T	-3.4475931	0.00118597	*SPy_1551*	hypothetical protein	ORF	S
473838	+	C	-2.6302029	0.04394521	*SPy_0587*	hypothetical protein	ORF	U
302525	+	C	-2.4917831	0.01146702	*murC*	UDP-N-acetylmuramate--L-alanine ligase	ORF	S

584148	+	G	-4.8362554	0.00265048	*ezrA*	septation ring formation regulator EzrA	ORF	U
278847	+	T	-3.3416019	0.00276394	*SPy_0316*	hypothetical protein	ORF	S
1154281	−	C	-3.5500251	0.0038051	*SPy_1391*	O-methyltransferase	ORF	S
1536397	−	C	-4.551466	0.0008995	*SPy_1851*	C3-degrading proteinase	ORF	U
836743	+	T	-2.73752	0.02099924	*acoB* ↔ *acoC*	"acetoin dehydrogenase (TPP-dependent) subunit beta" ↔ "branched-chain alpha-keto acid dehydrogenase E2 subunit"	UTR	U
1366859	+	C	-2.9196368	0.00568763	*recU*	Holliday junction-specific endonuclease	ORF	S
91733	+	T	-3.1455631	0.00609803	*pbp1b*	penicillin-binding protein 1b	ORF	S
285126	+	A	-3.5989427	0.00369495	*sstT*	serine/threonine transporter SstT	ORF	U
1152853	−	T	-2.3907009	0.01903583	Spy_sRNA1153040		UTR	U
71242	+	C	-4.4743989	0.00260679	*rpsH* ↔ *rplF*	"30S ribosomal protein S8" ↔ "50S ribosomal protein L6"	UTR	S
120819	+	T	-2.8742207	0.02117838	*SPy_0129*	hypothetical protein	ORF	U
1559098	−	T	-5.1115322	0.0002004	*glnA*	glutamine synthetase	ORF	U
849120	+	A	-3.6372278	0.00470206	*glmM*	phosphoglucosamine mutase	ORF	S
1556174	+	T	-3.8652108	0.00292	*rnjA*	Zn-dependent hydrolase	ORF	S
1432259	−	T	-4.0084674	0.00062484	*ecsB*	ABC transporter permease	ORF	U
706040	+	A	-3.7204913	0.00425789	*fruB*	fructose-1-phosphate kinase	ORF	S
115872	−	G	-3.865514	0.00120268	*rofA*	regulatory protein	ORF	U
502476	+	C	-3.0836874	0.00546513	*pacL*	calcium-transporting ATPase	ORF	S
1851196	+	G	-3.1728037	0.00901808	*htrA*	serine protease	ORF	U
409537	−	T	-2.7036679	0.00131067	*SPy_0508*	hypothetical protein	ORF	S
415383	+	T	-3.2051353	0.02869468	*dgs*	sugar transferase	ORF	U
275363	+	A	-3.7163707	0.01676677	*SPy_0309*	hypothetical protein	ORF	U
1558401	−	G	-3.4088237	0.00032401	*glnA*	glutamine synthetase	ORF	S
1454165	−	T	-2.6955879	0.00334994	*fabH*	3-oxoacyl-ACP synthase	ORF	U
666036	+	T	-3.1112758	0.00751673	*rlmI*	hypothetical protein	ORF	S
744602	+	T	-3.4176019	0.01563308	Spy_sRNA744487 ↔ *pyrF*	"orotidine 5'-phosphate decarboxylase"	UTR	U
1840324	−	A	-4.5140052	9.00E-05	*guaB*	inosine 5'-monophosphate dehydrogenase	ORF	U
88538	+	T	-4.4644751	0.00396	*adcB*	ABC transporter permease	ORF	U

1763860	−	T	-3.7898395	0.00268056	*SPy_2111*	hypothetical protein	ORF	U
411744	−	T	-4.5263088	0.00046146	*SPy_0512*	NAD(P)H-dependent quinone reductase	ORF	U
146883	+	T	-2.9185577	0.01567386	*purA*	adenylosuccinate synthetase	ORF	U
947611	+	C	-3.6198655	0.0137656	*gyrA*	DNA gyrase subunit A	ORF	U
931822	+	T	-3.0214728	0.02095553	*pbuX*	purine permease	ORF	U
640060	−	G	-4.2470927	0.00010443	*mscL*	large conductance mechanosensitive channel	ORF	S
1329559	+	T	-5.0813197	0.01418538	Spy_sRNA1329429 ↔ *SPy_1610*	"hypothetical protein"	UTR	U
1124061	−	C	-3.2889381	0.00383466	*murZ*	UDP-N-acetylglucosamine 1-carboxyvinyltransferase	ORF	U
51860	+	G	-2.9651687	0.02304978	*purB*	adenylosuccinate lyase	ORF	S
585163	+	G	-4.4310869	0.00405894	*ezrA*	septation ring formation regulator EzrA	ORF	U
1459561	−	G	-3.6226876	0.00674955	*grpE*	heat shock protein GrpE	ORF	U
1105854	−	T	-5.2398441	0.00130121	*obgE*	GTPase ObgE	ORF	U
1517716	+	C	-3.1120511	0.02142733	*corA*	divalent cation transport protein	ORF	U
1761325	−	T	-3.9648187	0.00040452	*nrdD*	anaerobic ribonucleoside triphosphate reductase	ORF	S
1565753	−	C	-4.1878428	0.00290846	*SPy_1885*	hypothetical protein	ORF	U
1266402	−	G	-3.2525311	0.00150099	*rsmD*	hypothetical protein	ORF	U
1057045	−	A	-3.4186853	0.01226775	*pyk*	pyruvate kinase	ORF	U
1109568	−	T	-3.4910506	0.00301534	*SPy_1340*	hypothetical protein	ORF	U
1026696	−	T	-2.9695957	0.01743404	*pstC2*	phosphate ABC transporter permease	ORF	U
758666	+	T	-2.6625223	0.0073664	*SPy_t40* ↔ *rpsA*	"tRNA-" ↔ "30S ribosomal protein S1"	UTR	U
351331	+	T	-3.6931896	0.01249206	*metS*	methionyl-tRNA synthetase	ORF	U
902708	+	T	-3.4278645	0.01619415	*potB*	spermidine / putrescine ABC transporter permease	ORF	U
1451322	−	A	-3.372407	0.00464364	*fabD*	ACP S-malonyltransferase	ORF	U
1014156	−	C	-3.2694341	0.01760568	*bmpA*	lipoprotein	ORF	U
1026888	−	T	-3.6408174	0.00648616	*pstC2*	phosphate ABC transporter permease	ORF	U
69496	+	G	-2.0300039	0.04394044	*rplN* ↔ *rplX*	"50S ribosomal protein L14" ↔ "50S ribosomal protein L24"	UTR	U
1453782	−	T	-5.6260181	0.00051279	*fabH*	3-oxoacyl-ACP synthase	ORF	U
185345	+	T	-2.9675033	0.01106193	*tgt*	queuine tRNA-ribosyltransferase	ORF	S

1611297	−	G	-4.0763607	0.00047754	*SPy_1936*	hypothetical protein	ORF	S
714002	+	T	-3.2876809	0.01000566	Spy_sRNA713953 ↔ *papS*	"tRNA CCA-pyrophosphorylase"	UTR	S
73158	+	A	-3.7312026	0.00747594	Spy_sRNA73113 ↔ *rplO*	"50S ribosomal protein L15"	UTR	U
266186	+	T	-2.9692707	0.00417597	*SPy_t35* ↔ *SPy_r08*	"tRNA-Ala (TGC)" ↔ "23S rRNA"	UTR	S
1124008	−	C	-2.7016039	0.04687215	*murZ*	UDP-N-acetylglucosamine 1-carboxyvinyltransferase	ORF	U
1698393	−	C	-5.9396749	6.58E-05	*SPy_2038*		ORF	U
18901	+	T	-3.0663442	0.0040905	*SPy_t01* ↔ *SPy_r02*	"23S rRNA"	UTR	S
24902	+	T	-2.985014	0.00510673	*SPy_t11* ↔ *SPy_r04*	"23S rRNA"	UTR	S
1225917	−	A	-3.1458554	0.00713146	*SPy_1494*	hypothetical protein	ORF	U
1581281	−	T	-2.7984486	0.00098071	*SPy_t55* ↔ *SPy_r11*	"tRNA-Ala (TGC)" ↔ "23S rRNA"	UTR	S
1334347	−	T	-3.1000049	0.00062425	*SPy_t50* ↔ *SPy_r09*	"tRNA-Ala ("↔ "23S rRNA"	UTR	S
81120	+	T	-2.8604307	0.00605576	*SPy_t22* ↔ *SPy_r06*	"tRNA-Ala (TGC)" ↔ "23S RNA"	UTR	S
319029	+	A	-4.1438573	0.00614776	*trmL*	23S rRNA methyltransferase	ORF	U
1443367	+	T	-1.9016982	0.02351018	Spy_sRNA1443306		UTR	S
1404844	−	A	-4.2445705	1.28E-05	Spy_sRNA1404921	*glyQ* leader	UTR	S
1053257	−	C	-5.3650373	0.00097458	*phnA*	hypothetical protein	ORF	U
1368291	+	T	-3.5666218	0.0096089	*pbp1A*	penicillin-binding protein 1a	ORF	U
72832	+	A	-3.6511853	0.00553136	*rpsE*	30S ribosomal protein S5	ORF	S
1698866	−	C	-4.8724699	0.00036038	*speB*	pyrogenic exotoxin B	ORF	U
185405	+	G	-3.4354811	0.02808139	*tgt*	queuine tRNA-ribosyltransferase	ORF	U
1455126	−	G	-3.1105306	0.01566637	*phaB*	enoyl-CoA hydratase	ORF	U
1441757	+	T	-2.5688785	0.0304393	*manM*	PTS system mannose-specific transporter subunit IIC	ORF	U
661518	+	C	-2.8453885	0.01791255	Spy_sRNA661400 ↔ *infC*	"translation initiation factor IF-3"	UTR	S
68255	+	G	-6.8473319	0.01594932	*rplP*	50S ribosomal protein L16	ORF	U
826106	+	T	-2.6039698	0.02021694	*Spy_sRNA825970*		UTR	U
1699898	−	T	-5.0563348	0.0027535	Spy_sRNA1699993 ↔ *speB*	"pyrogenic exotoxin B"	UTR	U
31491	+	A	-3.529689	0.00260679	*cdhA*	hypothetical protein	ORF	S
1295748	−	T	-3.2287988	0.0001929	Spy_sRNA1295825		UTR	S

1535220	−	T	-3.1389072	0.01831679	*SPy_1850* ↔ *pfl*	"esterase" ↔ "pyruvate formate-lyase"	UTR	U
1441617	+	T	-3.2547213	0.02864112	*manM*	PTS system mannose-specific transporter subunit IIC	ORF	U
639502	+	A	-3.1424387	0.0255539	*rpsU*	30S ribosomal protein S21	ORF	U
1683342	−	C	-3.8742051	0.00779785	*sic*	inhibitor of complement-mediated lysis	ORF	U
116813	+	A	-3.1096906	0.00611928	Spy_sRNA116781 *or cpa*		ORF	U
661502	+	C	-2.604117	0.01453765	Spy_sRNA661400		UTR	S
1465711	−	T	-3.4642669	0.03002366	*gatA*	aspartyl/glutamyl-tRNA amidotransferase subunit A	ORF	U
946419	−	T	-3.7279895	0.01193934	*ldh*	L-lactate dehydrogenase	ORF	U
615125	+	T	-3.1894959	0.01188897	*atpB*	ATP synthase F0F1 subunit A	ORF	U
1518451	+	C	-3.9067601	0.00297359	*SPy_1828*	hypothetical protein	ORF	U
1811863	−	A	-3.0695344	0.00783531	Spy_sRNA1811977 ↔ *rpsD*	"30S ribosomal protein S4"	UTR	U
1336230	−	G	-1.9729673	0.01159786	*yhbH* ↔ *SPy_r10*	"hypothetical protein" ↔ "16S rRNA"	UTR	U
17018	+	G	-1.9785249	0.01773377	*SPy_0016* ↔ *SPy_r01*	"amino acid permease" ↔ "16S rRNA"	UTR	U
79237	+	G	-2.0956978	0.01429475	Spy_sRNA78598 ↔ *SPy_r05*	"16S rRNA"	UTR	U
1137997	−	T	-3.6598094	0.00447574	*gapN*	NADP-dependent glyceraldehyde-3-phosphate dehydrogenase	ORF	U
1696694	+	C	-3.4968468	0.01259933	Spy_sRNA1696464 ↔ Spy_sRNA1696905		UTR	U
482849	−	G	-3.3230751	0.01063369	Spy_sRNA482963		UTR	U
31549	+	T	-2.4236401	0.02747926	*cdhA*	hypothetical protein	ORF	U
72163	+	G	-3.1657866	0.02850352	*rplR*	50S ribosomal protein L18	ORF	U

Table A4. RNA 3' ends more abundant in the WT strain than in Δ*rny* strain (3' *rny*_ends).
List of the RNA 3' ends identified as more abundant in the WT strain than in the Δ*rny* strain and named as 3' *rny*_ends (see Figure 18). The information is provided with regard the chromosome position, the directionality of the strand (positive (+) or negative (-)), the log 2 Fold Change (Log2FC), the False Discovery Rate (FDR), the gene ID (the symbol "↔" depicts the intergenic region between two genes), the description of the gene function based on NCBI, the location of the 3' *rny*_end (in open reading frame (ORF) and untranslated region (UTR)) and the type of 3' *rny*_end ("unique" (U) or "stepped" (S)). R.A.-B. generated the raw list of 3' *rny*_ends analysed by L.B., A.-L.L. and A.L.R.

	3' *rny*_end	Trimming stop position	Strand	Gene ID	Description	Location
PNPase	31491	31491	+	*cdhA*	Hypothetical protein	ORF
	31549	31553	+	*cdhA*	Hypothetical protein	ORF
	71242	71241	+	*rpsH* ↔ *rplF*	30S ribosomal protein S8 ↔ 50S ribosomal protein L6	UTR
	73158	73156	+	Spy_sRNA73113 ↔ *rplO*	50S ribosomal protein L15	UTR
	111312	111312	+	*ytpR*		ORF
	116813	116812	+	Spy_sRNA116781 *or cpa*	Collagen binding protein	ORF
	117634	117634	+	*cpa*	Collagen binding protein	ORF
	120819	120818	+	*SPy_0129*	Hypothetical protein	ORF
	185345	185344	+	*tgt*	Queuine tRNA-ribosyltransferase	ORF
	278847	278847	+	*SPy_0316*	Hypothetical protein	ORF
	287540	287540	−	*ntpJ*	V-type Na -ATPase subunit J	ORF
	377298	377298	+	*ftsK*	Hypothetical protein	ORF
	409537	409537	−	*SPy_0508*	Hypothetical protein	ORF
	477545	477545	−	*SPy_0593 (antisense)*	Hypothetical protein	UTR
	502476	502476	+	*pacL*	Calcium-transporting ATPase	ORF
	583794	583794	+	*gyrB* ↔ *ezrA*	DNA gyrase subunit B ↔ septation ring formation regulator EzrA	UTR
	615125	615125	+	*atpB*	ATP synthase F0F1 subunit A	ORF
	661502	661502	+	Spy_sRNA661400		UTR
	666036	666036	+	*rlmI*	Hypothetical protein	ORF
	706040	706040	+	*fruB*	Fructose-1-phosphate kinase	ORF
	758666	758666	+	*SPy_t40* ↔ *rpsA*	tRNA-Gln (TTG) ↔ 30S ribosomal protein S1	UTR
	836743	836743	+	*acoB* ↔ *acoC*	Acetoin dehydrogenase (TPP-dependent) subunit beta ↔ branched-chain alpha-keto acid dehydrogenase E2 subunit	UTR
	849120	849120	+	*glmM*	Phosphoglucosamine mutase	ORF
	861256	861256	+	Spy_sRNA860730 ↔ *lepA*	CRISPR RNA (crRNA) ↔ GTP-binding protein LepA	UTR
	902708	902711	+	*potB*	Spermidine / putrescine ABC transporter permease	ORF
	947611	947611	+	*gyrA*	DNA gyrase subunit A	ORF
	1014156	1014153	−	*bmpA*	Lipoprotein	ORF

	1023862	1023862	−	*pstB*	Phosphate transporter ATP-binding protein	ORF
	1057045	1057045	−	*pyk*	Pyruvate kinase	ORF
	1105854	1105854	−	*obgE*	GTPase ObgE	ORF
	1124061	1124061	−	*murZ*	UDP-N-acetylglucosamine 1-carboxyvinyltransferase	ORF
	1152119	1152119	−	*alaS*	Alanyl-tRNA synthetase	ORF
	1154281	1154280	−	*SPy_1391*	O-methyltransferase	ORF
	1225917	1225916	−	*SPy_1494*	Hypothetical protein	ORF
	1366859	1366859	+	*recU*	Holliday junction-specific endonuclease	ORF
	1401103	1401100	−	*glyS* ↔ *SPy_1687*	Glycyl-tRNA synthetase subunit beta ↔ hypothetical protein	UTR
	1404844	1404844	−	Spy_sRNA1404921	*glyQ* leader	UTR
	1441617	1441618	+	*manM*	PTS system mannose-specific transporter subunit IIC	ORF
	1454165	1454165	−	*fabH*	3-oxoacyl-ACP synthase	ORF
	1465711	1465711	−	*gatA*	Aspartyl/glutamyl-tRNA amidotransferase subunit A	ORF
	1518451	1518451	+	*SPy_1828*	Hypothetical protein	ORF
	1536397	1536397	−	*SPy_1851*	C3-degrading proteinase	ORF
	1556174	1556174	+	*rnjA*	Zn-dependent hydrolase	ORF
	1558401	1558401	−	*glnA*	Glutamine synthetase	ORF
	1811851	1811851	−	Spy_sRNA1811977 ↔ *rpsD*	30S ribosomal protein S4	UTR
	1840324	1840322	−	*guaB*	Inosine 5'-monophosphate dehydrogenase	ORF
YhaM	18901	18901	+	*SPy_t01* ↔ *SPy_r02*	23S rRNA	UTR
	24902	24902	+	*SPy_t11* ↔ *SPy_r04*	23S rRNA	UTR
	73158	73156	+	*Spy_sRNA73113* ↔ *rplO*	50S ribosomal protein L15	UTR
	81120	81120	+	*SPy_t22* ↔ *SPy_r06*	tRNA-Ala (TGC) ↔ 23S RNA	UTR
	116813	116812	+	Spy_sRNA116781 *or cpa*	Collagen binding protein	ORF
	266186	266186	+	*SPy_t35* ↔ *SPy_r08*	tRNA-Ala (TGC) ↔ 23S rRNA	UTR
	278847	278847	+	*SPy_0316*	Hypothetical protein	ORF
	292690	292690	+	*covS*	Sensory transduction histidine kinase	ORF
	302525	302522	+	*murC*	UDP-N-acetylmuramate--L-alanine ligase	ORF
	411744	411741	−	*SPy_0512*	NAD(P)H-dependent quinone reductase	ORF

	477545	477545	−	*SPy_0593* (antisense)	Hypothetical protein	UTR
	744602	744602	+	Spy_sRNA744487 ↔ *pyrF*	Orotidine 5'-phosphate decarboxylase	UTR
	826106	826106	+	Spy_sRNA825970		UTR
	861256	861256	+	Spy_sRNA860730 ↔ *lepA*	GTP-binding protein LepA	UTR
	1026696	1026697	−	*pstC2*	Phosphate ABC transporter permease	ORF
	1137997	1137997	−	*gapN*	NADP-dependent glyceraldehyde-3-phosphate dehydrogenase	ORF
	1152853	1152857	−	Spy_sRNA1153040		UTR
	1276420	1276418	−	*SPy_1551*	Hypothetical protein	ORF
	1334347	1334347	−	*SPy_t50* ↔ *SPy_r09*	tRNA-Ala (TGC) ↔ 23S rRNA	UTR
	1432259	1432259	−	*ecsB*	ABC transporter permease	ORF
	1454165	1454169	−	*fabH*	3-oxoacyl-ACP synthase	ORF
	1559098	1559098	−	*glnA*	Glutamine synthetase	ORF
	1581281	1581281	−	*SPy_t55* ↔ *SPy_r11*	tRNA-Ala (TGC) ↔ 23S rRNA	UTR
	1633958	1633958	−	*proS*	Prolyl-tRNA synthetase	ORF
	1761325	1761325	−	*nrdD*	Anaerobic ribonucleoside triphosphate reductase	ORF
	1762451	1762455	−	*nrdD*	Anaerobic ribonucleoside triphosphate reductase	ORF
	1763385	1763385	−	*SPy_2111*	Hypothetical protein	ORF
RNase R	1152853	1152853	−	Spy_sRNA1153040		UTR

Table A5. List of 3'*my*_ends that correspond to 3'-to-5' exoRNase trimming stop positions.
The 3' *my*_ends were compared to the trimming stop positions of PNPase, YhaM and RNase R (see Material & Methods and Figure 20A and B). The chromosome position of the matching 3' *my*_ends and trimming stops is reported. Information is also provided with regard the directionality of strand positive (+) or negative (-), the gene ID (the symbol "↔" depicts an intergenic region between two genes), the gene description (based on NCBI) and the location of the processing site (in untranslated region (UTR) and open reading frame (ORF)). The list of the 3' *my*_ends ccorresponding to 3'-to-5' exoRNase trimming stop positions was generated by T.T.R. and analyzed by L.B., A.-L.L. and A.L.R.

	3' *rny*_end	Trimming start position	Strand	Processivity	Gene ID	Location	nt	nt + 1	nt + 2	nt + 3
PNPase	31549	31714	+	165	*cdhA*	ORF	A	G	G	T
	71242	71295	+	53	*rpsH* ↔ *rplF*	UTR	G	T	G	A
	73158	73278	+	120	*rplO*	ORF	G	A	A	T
	116813	116859	+	46	Spy_sRNA116781 *or cpa*	ORF	T	C	A	A
	117634	117788	+	154	*cpa*	ORF	T	G	G	A
	120819	121011	+	192	*SPy_0129*	ORF	C	G	T	A
	278847	278883	+	36	*SPy_0316*	ORF	G	G	A	T
	409537	409513	−	24	*SPy_0508*	ORF	T	C	G	A
	417307	417329	+	22	Spy_sRNA416968 ↔ *thrS*	UTR	T	G	A	C
	477545	477426	−	119	*SPy_0593* ↔ *SPy_0591*	UTR	T	C	G	A
	615125	615216	+	91	*atpB*	ORF	T	G	C	T
	661502	661554	+	52	*infC*	ORF	G	G	A	T
	758666	758684	+	18	*rpsA*	ORF	T	T	G	A
	1014156	1014144	−	12	*bmpA*	ORF	G	A	T	T
	1023862	1023727	−	135	*phoU*	ORF	G	A	A	A
	1401103	1401097	−	6	*SPy_1687*	ORF	G	G	A	T
	1404844	1404774	−	70	Spy_sRNA1404921	UTR	T	G	A	A
	1441617	1441793	+	176	*manN*	ORF	G	A	A	C
	1811851	1811826	−	25	Spy_sRNA1811977 ↔ *rpsD*	UTR	G	T	A	A
YhaM	278847	278849	+	2	*SPy_0316*	ORF	A	T	G	C
	292690	292696	+	6	*covS*	ORF	G	T	A	A
	302525	302528	+	3	*murC*	ORF	G	G	A	C
	1762451	1762450	−	1	*nrdD*	ORF	G	A	T	A
	411744	411734	−	10	*SPy_0512*	ORF	G	G	T	T

Table A6. List of the initial RNase Y processing sites identified as 3'-to-5' exoRNase trimming start positions.
The initial RNase Y processing positions were identified by searching for 3'-to-5' exoRNase trimming start positions downstream of the 3' *rny*_ends corresponding to trimming stop positions (in Table A5), see Material & Methods and Figure 20B and C. The chromosome position of the 3' *rny*_ends and of the trimming start positions

(corresponding to the initial RNase Y processing site) is reported. Information is also provided with regard the directionality of the strand (positive (+) or negative (-)), the processivity (*i.e.* distance between the two positions), the gene ID (the symbol "↔" depicts an intergenic region between two genes), the gene description (based on NCBI) and the location of the processing site (in untranslated region (UTR) and open reading frame (ORF)) and the nucleotide (nt) at the 3'-to-5' exoRNase trimming start position and also the nt at the 3 following positions. The list of the 3'-to-5' exoRNase trimming start positions located downstream of the 3' *rny*_ends corresponding to 3'-to-5' exoRNase trimming stop positions was generated by T.T.R. and analyzed by L.B., A.-L.L. and A.L.R.

	3′ *rny*_end	Trimming start position	Strand	Distance	Gene ID	Location	nt
PNPase	115872	115872	−	0	*rofA*	ORF	G
	482849	482849	−	0	Spy_sRNA482963	UTR	G
	585163	585163	+	0	*ezrA*	ORF	G
	639502	639506	+	4	*rpsU*	ORF	G
	1696694	1696695	+	1	Spy_sRNA1696464 ↔ Spy_sRNA1696905	UTR	G
	1851196	1851196	+	0	*htrA*	ORF	G
YhaM	1455126	1455125	−	1	*phaB*	ORF	A
	1405169	1405167	−	2	*SPy_1691*	ORF	A

Table A7. List of 3'*rny*_ends that correspond to 3'-to-5' exoRNase trimming start positions.
The 3' *rny*_ends were compared to the trimming start positions of PNPase, YhaM and RNase R (see Material & Methods and Figure 20A and D). The chromosome position of the matching 3' *rny*_ends and trimming starts is reported. Information is also provided with regard the directionality of strand (positive (+) or negative (-)), the distance between the 3' *rny*_end and the trimming start positions, the gene ID (the symbol "↔" depicts an intergenic region between two genes), the gene description (based on NCBI) and the location of the processing site (in untranslated region (UTR) and open reading frame (ORF)) and the nucleotide (nt) at the 3'-to-5' exoRNase trimming start position. The list of the 3' *rny*_ends ccorresponding to 3'-to-5' exoRNase trimming start positions was generated by T.T.R. and analyzed by L.B., A.-L.L. and A.L.R.

5′ *rny*_end	3′ *rny*_end	Distance	Strand	Gene ID	Location	Trimming stop position		Trimming start position	
Fragments trimmed by PNPase									
31422	31491	69	+	*cdhA*	ORF	31491			
71149	71242	93	+	*rpsH* ↔ *rplF*	UTR	71241		71295	
111220	111312	92	+	*ytpR*	ORF	111312			
117541	117634	93	+	*cpa*	ORF	117634			
120684	120819	135	+	*SPy_0129*	ORF	120818			
287634	287540	94	−	*ntpJ*	ORF	287540			
409647	409537	110	−	*SPy_0508*	ORF	409537		409513	
614992	615125	133	+	*atpB*	ORF	615125		615216	
665855	666036	181	+	*rlmI*	ORF	666036			
705947	706040	93	+	*fruB*	ORF	706040			
849024	849120	96	+	*glmM*	ORF	849120			
947461	947611	150	+	*gyrA*	ORF	947611			
1023988	1023862	126	−	*pstB*	ORF	1023862			
1152301	1152119	182	−	*alaS*	ORF	1152119			
1366777	1366859	82	+	*recU*	ORF	1366859			
1401255	1401103	152	−	*glyS*	ORF	1401100		1401097	
1518318	1518451	133	+	*SPy_1828*	ORF	1518451			
1558584	1558401	183	−	*glnA*	ORF	1558401			
1840448	1840324	124	−	*guaB*	ORF	1840322			
Fragments trimmed by PNPase and YhaM									
						PNPase	**YhaM**	**PNPase**	**YhaM**
73054	73158	104	+	*rpmD* ↔ Spy_sRNA73113	UTR	73156	73156	73278	73158
278753	278847	94	+	*SPy_0316*	ORF	278847	278847	278883	278849
1454339	1454165	174	−	*fabH*	ORF	1454165	1454169		
Fragments trimmed by YhaM									
302413	302525	112	+	*murC*	ORF	302522		302528	
744449	744602	153	+	*SPy_0899* ↔ SpRNA744487	UTR	744602			

1152937	1152853	84	−	Spy_sRNA1153040	UTR	1152857	1152854
1276533	1276420	113	−	*SPy_1551*	ORF	1276418	
1634045	1633958	87	−	*proS*	ORF	1633958	
1761437	1761325	112	−	*nrdD*	ORF	1761325	
1762537	1762451	86	−	*SPy_2111* ↔ *nrdD*	UTR	1762455	1762450
1763486	1763385	101	−	*SPy_2111*	ORF	1763385	
292611	292690	79	+	*covS*	ORF	292690	292696
Fragments not identified as trimmed							
88476	88538	62	+	*adcB*	ORF		
1698455	1698393	62	−	*SPy_2038*	ORF		
69330	69496	166	+	*rplN*	ORF		
115985	115872	113	−	*rofA*	ORF		
146735	146883	148	+	*purA*	ORF		
275237	275363	126	+	*SPy_0309*	ORF		
285038	285126	88	+	*sstT*	ORF		
318951	319029	78	+	*trmL*	ORF		
323093	323184	91	+	*hlyX*	ORF		
381172	381252	80	+	*frr* ↔ *SPy_0464*	UTR		
473749	473838	89	+	*SPy_0587*	ORF		
713911	714002	91	+	*SPy_0865* ↔ Spy_sRNA713953	UTR		
1053328	1053257	71	−	*phnA*	ORF		
1152840	1152709	131	−	*Spy_sRNA1153040*	UTR		
1329406	1329559	153	+	*SPy_1608* ↔ Spy_sRNA1329429	UTR		
1453867	1453782	85	−	*fabH*	ORF		
1455284	1455126	158	−	*phaB*	ORF		
1517566	1517716	150	+	*corA*	ORF		
1535306	1535220	86	−	*SPy_1850*	ORF		
1611380	1611297	83	−	*SPy_1936*	ORF		
1764034	1763860	174	−	*SPy_2111*	ORF		

Table A8. List of short RNA fragments generated by RNase Y processing.
The chromosome positions of the 3' *rny*_ends and 5' *rny*_ends that were found from 50 to 200 nt of distance from each other are reported (see Figure 24). The information is provided with regard the distance in nt between the 3' *rny*_ends and 5' *rny*_ends, the directionality of the strand (positive (+) or negative (-)), the gene ID (the symbol "↔" depicts an intergenic region between two genes), the gene description (based on NCBI) and the location of the processing site (in untranslated region (UTR) and open reading frame (ORF)). The list of the paired 3' *rny*_ends and 5' *rny*_ends was generated by T.T.R. and analyzed by L.B., A.-L.L. and A.L.R.

	5' *rny*_end	Trimming start position	Strand	Distance	Gene ID	Description	Location
PNPase	14744	14743	+	1	*ftsH*	cell division protein	ORF
	662728	662727	+	1	*rplT* ↔ *ltaS*	"50S ribosomal protein L20" ↔ "hypothetical protein"	UTR
	1625568	1625569	−	1	*rpsO*	30S ribosomal protein S15	ORF
	1844913	1844912	+	1	*SPy_2210*	ABC transporter ATP-binding protein	ORF
	897129	897127	+	2	*folC.1*	folyl-polyglutamate synthetase	ORF
	1057924	1057926	−	2	*pfkA* ↔ *pyk*	"6-phosphofructokinase" ↔ "pyruvate kinase"	UTR
	1353157	1353159	−	2	*rpoZ*	DNA-directed RNA polymerase subunit omega	ORF
	1400285	1400287	−	2	*SPy_1686* ↔ Spy_sRNA1400292	"hypothetical protein"	UTR
	1518318	1518316	+	2	*SPy_1828*	hypothetical protein	ORF
	1530955	1530953	+	2	*SPy_1845*	hypothetical protein	ORF
	1764034	1764036	−	2	*SPy_2111*	hypothetical protein	ORF
	73054	73051	+	3	*rpmD*	50S ribosomal protein L30	UTR
	1014753	1014756	−	3	*cdd* ↔ *bmpA*	"cytidine deaminase" ↔ "lipoprotein"	UTR
	1455284	1455287	−	3	*phaB*	enoyl-CoA hydratase	ORF
	1387143	1387147	−	4	*SPy_1673*	ABC transporter permease	ORF
	1430377	1430381	−	4	*trmB* ↔ *SPy_t51*	"tRNA (guanine-N(7)-)-methyltransferase" ↔ "tRNA-Ser (GGA)"	UTR
	849003	848998	+	5	*glmM*	phosphoglucosamine mutase	ORF
	1404769	1404774	−	5	Spy_sRNA1404921		UTR
	71140	71134	+	6	*rpsH*	30S ribosomal protein S8	ORF
	1764633	1764640	−	7	*SPy_2112*	hypothetical protein	ORF
	78458	78450	+	8	*rplQ*	50S ribosomal protein L17	ORF
	1761437	1761446	−	9	*nrdD*	anaerobic ribonucleoside triphosphate reductase	ORF
	287642	287652	−	10	*ntpJ*	V-type Na-ATPase subunit J	ORF

Table A9. List of 5'*rny*_ends paired to PNPase trimming start positions.
The 5' *rny*_ends were associated to the trimming start positions of PNPase, YhaM and RNase R (see Material & Methods and Figure 26C). The chromosome position of the paired 5' *rny*_ends and trimming starts is reported. The 5' *rny*_ends were paired only to PNPase trimming start positions. Information is also provided with the regard the directionality of strand (positive (+) or negative (-)), the distance between the 5' *rny*_end and the PNPase trimming start positions, the gene ID (the symbol "↔"

depicts an intergenic region between two genes), the gene description (based on NCBI) and the location of the processing site (in untranslated region (UTR) and open reading frame (ORF)). The list of the 5' *my*_ends paired to PNPase start positions was generated by T.T.R. and analyzed by L.B., A.-L.L. and A.L.R.

Position	Strand	End	nt in 3'	Log 2 FC	FDR	Gene ID	Description	Location	Type
1412318	−	5	A	3.244795993	0.00261103	*SPy_1700*	hypothetical protein	ORF	U
31777	+	3	T	4.864234462	0.001395021	*cdhA* ↔ *prsA*	"hypothetical protein" ↔ "ribose-phosphate pyrophosphokinase"	UTR	U
1186492	+	5	A	4.450264832	0.001176324	*pyrD*	dihydroorotate dehydrogenase 1A	ORF	U
1547173	−	5	G	3.162110058	0.007856842	*SPy_1865*	hypothetical protein	ORF	U
1130742	+	5	G	3.893188339	0.001540553	*SPy_1363* (antisense)	hypothetical protein	UTR	S
727437	+	5	A	3.219844174	0.002063885	*idi* ↔ *mvaS_1*	"isopentenyl pyrophosphate isomerase" ↔ "3-hydroxy-3-methylglutaryl-coenzyme A"	UTR	S
279077	+	5	A	4.722613683	0.001176324	*SPy_0316*	hypothetical protein	ORF	U
75054	+	3	T	2.844357304	0.006147763	*secY* ↔ *adk*	"preprotein translocase subunit SecY" ↔ "adenylate kinase"	UTR	U
148549	+	3	T	2.993031075	0.004034073	*SPy_0163* ↔ *nusG*	"ABC transporter lipoprotein" ↔ "transcription antitermination protein NusG"	UTR	U
929283	+	5	G	3.384308457	0.003315777	*opuC*	ABC transporter binding protein	ORF	U
1528152	+	5	G	3.931316635	0.001540553	*spi*	signal peptidase I	ORF	S
1609322	−	5	G	4.006491275	0.001684618	*rpsI*	30S ribosomal protein S9	ORF	U
279187	+	3	A	3.440779327	0.013503824	*SPy_0316* ↔ *SPy_0317*	"hypothetical protein" ↔ "ABC transporter substrate-binding protein"	UTR	S
1330497	−	5	A	3.483996284	0.000597033	*SPy_t43*	tRNA-Glu (TTC)	UTR	U
713953	+	5	A	5.210396624	0.001487154	Spy_sRNA713953		UTR	S
115179	−	5	A	3.339084724	0.003709849	*rofA*	regulatory protein	ORF	U
292739	+	3	T	4.122776922	0.00091425	*covS*	sensory transduction histidine kinase	ORF	S
319134	+	3	A	3.134577129	0.005191423	*trmL* ↔ Spy_sRNA319190	"23S rRNA methyltransferase"	UTR	U
320688	+	3	T	3.637082831	0.004208683	Spy_sRNA320648		UTR	U
1108344	−	5	A	5.191850828	0.010839866	*rsuA*	16S pseudouridylate synthase	ORF	U
292628	+	5	A	3.110889408	0.002229995	*covS*	sensory transduction histidine kinase	ORF	U
579705	+	5	G	5.890580428	0.001176324	Spy_sRNA579683		UTR	U
338836	+	3	T	3.825351226	0.007366398	*SPy_0405*	DNA replication intiation control protein YabA	ORF	U
419327	+	3	T	3.760282357	0.004034073	*thrS* ↔ *SPy_0518*	"threonyl-tRNA synthetase" ↔ "ABC transporter ATP-binding protein"	UTR	S
477312	+	3	T	4.190742522	0.001972553	*SPy_0591* ↔ *SPy_0593*	"protease" ↔ "hypothetical protein"	UTR	U
580158	+	3	A	4.848729018	9.47E-05	*rplS* ↔ *SPy_t38*	"50S ribosomal protein L19" ↔ "tRNA-Arg (CCT)"	UTR	S

1746369	+	5	C	3.761395227	0.001176324	*SPy_2091* ↔ *rpsB*	"regulatory protein" ↔ "30S ribosomal protein S2"	UTR	U
588004	+	3	T	3.269186155	0.002288374	*eno* ↔ *SPy_0732*	"phosphopyruvate hydratase" ↔ "hypothetical protein"	UTR	S
714055	+	3	T	4.437644172	0.001188554	*papS*	tRNA CCA-pyrophosphorylase	ORF	S
1144148	+	5	A	3.786668655	0.001540553	*nrdE.2* ↔ *nrdF2*	"ribonucleotide-diphosphate reductase subunit alpha" ↔ "ribonucleotide-diphosphate reductase subunit beta"	UTR	S
1838938	−	5	A	4.932955795	0.003446201	Spy_sRNA1839037 ↔ *glcU*	"hypothetical protein"	UTR	U
639523	+	5	A	2.863512655	0.005361114	*rpsU*	30S ribosomal protein S21	ORF	S
727562	+	3	T	7.562067715	0.002606791	*idi* ↔ *mvaS_1*	"isopentenyl pyrophosphate isomerase" ↔ "3-hydroxy-3-methylglutaryl-coenzyme A"	UTR	S
826149	+	3	T	3.294671728	0.004090498	Spy_sRNA825970 ↔ *SPy_1016*	"hypothetical protein"	UTR	U
1821630	−	5	A	2.999134765	0.027237832	*mnmA*	tRNA-specific 2-thiouridylase MnmA	ORF	U
834702	+	3	C	3.000933933	0.035924422	*acoA*	acetoin dehydrogenase (TPP-dependent) subunit alpha	ORF	U
900380	+	3	T	3.419214271	0.015731497	*folK* ↔ *murB*	"2-amino-4-hydroxy-6- hydroxymethyldihydropteridine pyrophosphokinase" ↔ "UDP-N-acetylenolpyruvoylglucosamine reductase"	UTR	S
319526	+	5	T	3.342980327	0.028051389	*ribU*	hypothetical protein	ORF	U
1108243	−	5	G	2.864832125	0.005191277	*rsuA*	16S pseudouridylate synthase	ORF	U
864784	+	5	A	3.227903282	0.00367721	*sciB*	hypothetical protein	ORF	U
1065236	+	3	C	2.146581975	0.027637579	*SPy_s01 or* Spy_sRNA1065030	"tmRNA"	ORF	U
799582	+	5	C	3.442340784	0.006937556	*SPy_0976*	hypothetical protein	ORF	S
1633069	−	5	A	4.141392793	0.002021061	*proS*	prolyl-tRNA synthetase	ORF	U
1115389	−	3	T	3.851533941	0.005745952	*psr* ↔ *rumA*	"PBP 5 synthesis repressor" ↔ "RNA methyltransferase"	UTR	S
1053815	−	5	G	2.826755011	0.033393607	*glmS*	glucosamine--fructose-6-phosphate aminotransferase	ORF	U
279091	+	5	A	4.814114689	0.001540553	*SPy_0316*	hypothetical protein	ORF	U
1122830	−	3	T	4.807025719	3.53E-05	*grab* ↔ *rimJ*	"protein G-like alpha 2M-binding protein" ↔ "acetyl transferase"	UTR	S
917331	−	5	A	2.440070502	0.016897091	*SPy_1119*	hypothetical protein	ORF	U
1696503	+	5	T	4.729383544	0.001176324	*BILSSTRPYOSF370V 1_01732T*		ORF	U
1130843	+	3	G	4.237316485	0.001972553	*SPy_1363* ↔ *dnaX*	"hypothetical protein" ↔ "DNA polymerase III subunits gamma and tau"	UTR	S
759365	+	5	G	3.507232645	0.007214717	*rpsA*	30S ribosomal protein S1	ORF	U
985204	−	5	A	3.221124257	0.018937515	*SPy_1202*	GntR family transcriptional regulator	ORF	U

1108207	−	5	C	2.577212306	0.00688493	*rsuA*	16S pseudouridylate synthase	ORF	U
1144343	+	3	T	3.175542327	0.002972542	*nrdE.2* ↔ *nrdF2*	"ribonucleotide-diphosphate reductase subunit alpha" ↔ "ribonucleotide-diphosphate reductase subunit beta"	UTR	U
1839292	−	5	G	3.876062579	0.002426336	*guaB*	inosine 5'-monophosphate dehydrogenase	ORF	U
1152858	−	3	T	3.921362055	2.26E-05	Spy_sRNA1153040		UTR	S
866502	+	5	C	2.951959969	0.00902615	*SPy_1056*	hypothetical protein	ORF	U
759781	+	5	C	2.653863581	0.004502182	*rpsA*	30S ribosomal protein S1	ORF	U
1722836	−	5	T	3.429178026	0.009340337	*groEL*	molecular chaperone GroEL	ORF	U
1186671	+	3	T	3.494584094	0.002200297	*pyrD* ↔ Spy_sRNA1186876	"dihydroorotate dehydrogenase 1A"	UTR	S
764039	−	5	G	2.947438158	0.020263336	*SPy_0919*	hypothetical protein	ORF	U
1144105	+	5	A	3.891549698	0.002491144	*nrdE.2* ↔ *nrdF2*	"ribonucleotide-diphosphate reductase subunit alpha" ↔ "ribonucleotide-diphosphate reductase subunit beta"	UTR	U
1747099	+	5	A	4.664244977	0.002159742	*rpsB*	30S ribosomal protein S2	ORF	U
1747552	+	5	C	4.444347355	0.001540553	*tsf*	elongation factor Ts	ORF	S
1330352	−	3	T	1.998032027	0.018905098	Spy_sRNA1330456 ↔ *SPy_1610*	"hypothetical protein"	UTR	U
1362429	−	3	T	3.46021708	0.000334488	*mapZ* ↔ *luxS*	"hypothetical protein" ↔ "S-ribosylhomocysteinase"	UTR	S
1747452	+	5	C	6.070343159	0.001176324	*tsf*	elongation factor Ts	ORF	U
1412225	−	3	A	3.601123157	0.002254547	*SPy_1699* (antisense)	transcriptional regulator	UTR	S
1746402	+	5	C	4.570385015	0.001176324	*SPy_2091* ↔ *rpsB*	"regulatory protein" ↔ "30S ribosomal protein S2"	UTR	U
1547061	−	3	T	5.39976177	0.000155917	*dnaQ*	DNA polymerase III subunit epsilon	ORF	U
1564217	−	3	T	4.157821937	0.001494411	*SPy_1885* ↔ *SPy_1884*	"hypothetical protein" ↔ "several hypersensitive-induced response proteins"	UTR	U
1746359	+	5	C	3.708363188	0.001540553	*SPy_2091* ↔ *rpsB*	"regulatory protein" ↔ "30S ribosomal protein S2"	UTR	U
1632984	−	3	T	4.417898184	0.000290432	*proS* ↔ *polC*	"prolyl-tRNA synthetase" ↔ "DNA polymerase III PolC"	UTR	U
1820651	−	5	T	4.551197459	0.011935722	*mnmA* ↔ Spy_sRNA1820593	"tRNA-specific 2-thiouridylase MnmA"	UTR	U
587903	+	5	G	3.669307524	0.001540553	*eno* ↔ *SPy_0732*	"phosphopyruvate hydratase" ↔ "hypothetical protein"	UTR	S
1611214	−	5	C	2.673843167	0.007081647	*SPy_1936*	hypothetical protein	ORF	U
1365697	−	5	C	4.812691898	0.004926863	*gpsB*	hypothetical protein	ORF	U
1696783	−	3	C	2.969239837	0.000334488	Spy_sRNA1696905 ↔ Spy_sRNA1696464		UTR	U

1115544	−	5	G	4.674251651	0.005409722	*psr*	PBP 5 synthesis repressor	ORF	U
579750	+	5	A	4.150242413	0.001540553	Spy_sRNA579683 ↔ *rplS*	"50S ribosomal protein L19"	UTR	U
1701760	−	3	T	4.194774524	0.000467351	*mf* ↔ *ropB*	"mitogenic factor" ↔ "transcriptional regulator"	UTR	S
1747070	+	3	A	5.557468515	0.002066464	*rpsB*	30S ribosomal protein S2	ORF	U
1747315	+	3	T	4.597213274	0.000149841	*rpsB* ↔ *tsf*	"30S ribosomal protein S2" ↔ "elongation factor Ts"	UTR	S
1365865	−	5	T	3.588992733	0.004830991	*gpsB*	hypothetical protein	ORF	U
1747475	+	5	T	6.677236028	0.001176324	*tsf*	elongation factor Ts	ORF	S
1747474	+	3	T	4.283558874	0.001045368	*tsf*	elongation factor Ts	ORF	S
1747596	+	3	T	5.59603246	0.002650485	*tsf*	elongation factor Ts	ORF	U
1365841	−	5	C	4.423587162	0.003390621	*gpsB*	hypothetical protein	ORF	U
378719	+	5	G	3.035908755	0.00462229	*rplK* ↔ *rplA*	"50S ribosomal protein L11" ↔ "50S ribosomal protein L1"	UTR	U
1747690	+	3	C	5.444698369	0.002650485	*tsf*	elongation factor Ts	ORF	U
1053640	−	5	G	2.888051895	0.005968461	*glmS*	glucosamine--fructose-6-phosphate aminotransferase	ORF	U
1341253	+	5	T	3.954196474	0.011653679	*cysM*	O-acetylserine lyase	ORF	U
1519103	−	5	T	4.042479513	0.00908257	*ssb*	single-stranded DNA-binding protein	ORF	U
1748450	+	3	T	4.706143711	0.000149841	*tsf* ↔ *pepO*	"elongation factor Ts" ↔ "endopeptidase O"	UTR	S
1774730	+	3	T	3.448833775	0.00522187	*int4* ↔ Spy_sRNA1774740	"integrase"	UTR	U
1774768	+	3	T	3.262800682	0.002200297	Spy_sRNA1774740		UTR	S
577834	+	5	C	3.09750458	0.036116287	*SPy_0721*	flavodoxin	ORF	U
1747337	+	5	T	4.540070372	0.002472901	*rpsB* ↔ *tsf*	"30S ribosomal protein S2" ↔ "elongation factor Ts"	UTR	S
1824525	−	3	T	3.060858593	0.023350155	*SPy_2191* ↔ *sdhA*	"hypothetical protein" ↔ "L-serine dehydratase subunit alpha"	UTR	U
1404830	−	5	T	4.122351516	0.000597033	Spy_sRNA1404921	*glyQ* leader	UTR	S
1747787	+	5	T	6.275497155	0.001455677	*tsf*	elongation factor Ts	ORF	U

Table A10: Transcript ends (5' or 3') more abundant in the Δ*rny* strain than in the WT strain.
List of the RNA ends identified as more abundant in the Δ*rny* strain than in the WT strain, named as 5' or 3' Δ*rny*_ends. The information is provided with regard the chromosome position, the nature of the end (5' or 3'), the directionality of the strand (positive (+) or negative (-)), the log 2 Fold Change (Log2FC), the False Discovery Rate (FDR), the gene ID (the arrow "↔"depicts the intergenic region between two genes), description of the gene function based on NCBI, the location of the Δ*rny*_end

(in open reading frame (ORF) and untranslated region (UTR)) and the type of end ("unique" (U) or "stepped" (S)). R.A.-B. generated the raw list of Δ*rny*_ends analysed by L.B, A.-L.L. and A.L.R.

6.2 Declaration of independent work

I, Laura Broglia declare that I have completed the submitted thesis, entitled "*Regulating with ribonucleases in* Streptococcus pyogenes", independently using only the aids and tools specified.

I declared that the data presented in this work derived from my own original research. When the data were jointly generated with other researchers, I declare that I have, without exception, stated the respective contribution.

I expressly declare that all sources used in the abovementioned work have been properly referenced as such.

I am aware that violations against the principles of academic independence are considered deception and are punished accordingly.

Date Signature

6.3 List of publications

1) Broglia L†., Lécrivain A.L.†, Renault T.T., Hahnke K., Ahmed-Begrich R., Le Rhun A.*, and Emmanuelle Charpentier*. (2019)
An RNA-seq based comparative approach unravels the transcriptome-wide interplay between 3' to-5' exoRNases and RNase Y in *Streptococcus pyogenes*. In revision in *Nat. Commun*.
†co-first authors; *co-corresponding authors

2) Broglia L., Materne S., Lécrivain A.L., Hahnke K., Le Rhun A*., and Charpentier E*. (2018)
RNase Y-mediated regulation of the streptococcal pyrogenic exotoxin B. (2018) *RNA Biol.*, **15**, 1336–1347.
*co-corresponding authors

3) Le Rhun A†., Lécrivain A.L.†, Reimegård J., Proux-Wéra E., **Broglia L.**, Della Beffa C., and Charpentier E. (2017)
Identification of endoribonuclease specific cleavage positions reveals novel targets of RNase III in *Streptococcus pyogenes*. *Nucleic Acids Res.*, **45**, 2329–2340.
†co-first authors